

令和元年度 修士論文

聴覚皮質の要素的な音成分の結合における  
脳リズムの役割

学籍番号 1833054

氏名 小坂奏平

基盤理工学専攻

指導教員 檜森与志喜 教授

提出日 令和2年1月23日



## 目次

1. 序論 .....	3
1.1 はじめに .....	3
1.2 本研究における学会発表 .....	4
2. 基礎知識 .....	5
2.1 ソナグラム .....	5
2.2 音節について .....	6
2.3 音の受容器 .....	8
2.4 聴覚経路 .....	10
3. 研究背景 .....	11
3.1 一次聴覚野 .....	11
3.2 音情報のカテゴリー化 .....	18
3.3 言語処理と脳内リズム .....	21
4. 本研究の問題 .....	25
4.1 本研究の問題 .....	25
4.2 アプローチ .....	25

5. モデル.....	26
5.1 モデル概要 .....	26
5.2 一次聴覚野 (A1 層) .....	27
5.3 Feature-detective layer (FD 層) .....	29
5.4 Feature-binding layer (FB 層) .....	31
6. 結果 .....	35
6.1 A1 層の単一シラブル音に対する応答.....	35
6.2 FD 層のシラブル音に対する応答 .....	37
6.3 FB 層におけるシラブル内の子音—母音要素の結合 .....	43
6.4 連続したシラブルに対する各部位の応答 .....	47
7. 結論 .....	50
8. 今後の課題.....	50
9. 謝辞 .....	51
10. 参考文献 .....	52
11. 付録 .....	53

## 1. 序論

### 1.1 はじめに

聴覚による外界の認識は、生物にとって非常に重要な機能の一つである。夜行性の動物であるコウモリなどは、エコーロケーションによって標的までの距離や大きさ、その他森などの背景といった周囲の状況を知ることができる。また、夜行性でない我々人類や鳥類などもコミュニケーションの手段として聴覚を活用している。このように、様々な生物が聴覚を用いて自然環境に適応している。

このような聴覚による認識の神経機構を明らかにするために神経生理学、解剖学、心理学、行動学的な研究が行われてきた。しかし、聴覚は視覚に関する研究に比べて、まだ明らかになっていない点が多い。その理由は、一般的に聴覚系の神経細胞が視覚系の神経細胞より小さく、インパルス活動の計測が困難であること、また、静的な刺激を用いての研究が進んできた視覚系の研究に比べて、聴覚系では刺激が常に時間的に変動するため解析が難しい点にある。このように、聴覚系において認識のメカニズムは不明な点が多く、特に高次中枢での情報処理についてはあまり研究が進んでおらず、その知見も少ない。

現在、高次中枢での聴覚の研究は一次聴覚野が主である。一次聴覚野の光学測定において、音刺激に対するニューロンの活動は静的なものでなく、様々な周波数領域にわたって動的に変化するものであり、神経活動の時空間パターンとして表現されることが示されている。また、脳波の測定などから脳内には神経活動の振動（脳リズム）が存在し、聴覚系においても音刺激の認知に貢献する可能性が示唆されている。しかし、一次聴覚野での神経活動の時空間パターンからどのように音情報が表現され認知に至るか、また、脳リズムが認知にどのように関与しているかはわかっていない。

本研究では、これらの問題を解決するため、一次聴覚野を含む聴覚系のニューラルネットワークを作成し、音情報が認知されるメカニズム、脳リズムの効果について調べた。

本研究で作成したモデルは、一次聴覚野（A1）層、feature-detective（FD）層、feature-binding（FB）層の3層から構成される。A1層では、音声刺激を神経活動の時空間的なパターンに変換する。FD層では、A1層のパターンから音情報の特徴（子音、母音など）を抽出する。FB層では、FD層の情報をまとめ上げ、音節ごとに結合する。

本研究の結果は、一時聴覚野での神経活動の時空間パターンから音節情報をどのように取り出しているか、また、音情報の認知に脳リズムがどのように貢献しているかについて新しい知見を与えるものである。

## 1.2 本研究における学会発表

### 1.2.1

著者 : Sohei Kosaka, Yoshiki Kashimori  
題目 : A role of the brain rhythms in binding of elementary sound  
components in auditory cortex  
公表 : 第 42 回日本神経科学大会  
日時 : 2019 年 7 月 25 日

### 1.2.2

著者 : Sohei Kosaka, Yoshiki Kashimori  
題目 : Binding of elementary sound components by beta rhythms in  
auditory cortex  
公表 : 日本神経回路学会第 29 回全国大会 (JNNS2019)  
日時 : 2019 年 9 月 4 日

## 2. 基礎知識

### 2.1 ソナグラム

音の物理的な性質は、周波数、振幅（音圧）および時間の3つの要素から構成される。一般に用いられているソナグラムでは、周波数と時間をそれぞれ縦軸と横軸にとり、各周波数成分の振幅を線の濃淡で表す（図 2.1-1）。ソナグラムによって、音刺激にどのような周波数成分がどれくらい含まれるかが視覚的にわかる。音の時間的変動は、このように、周波数成分の時間的変化として表現される。

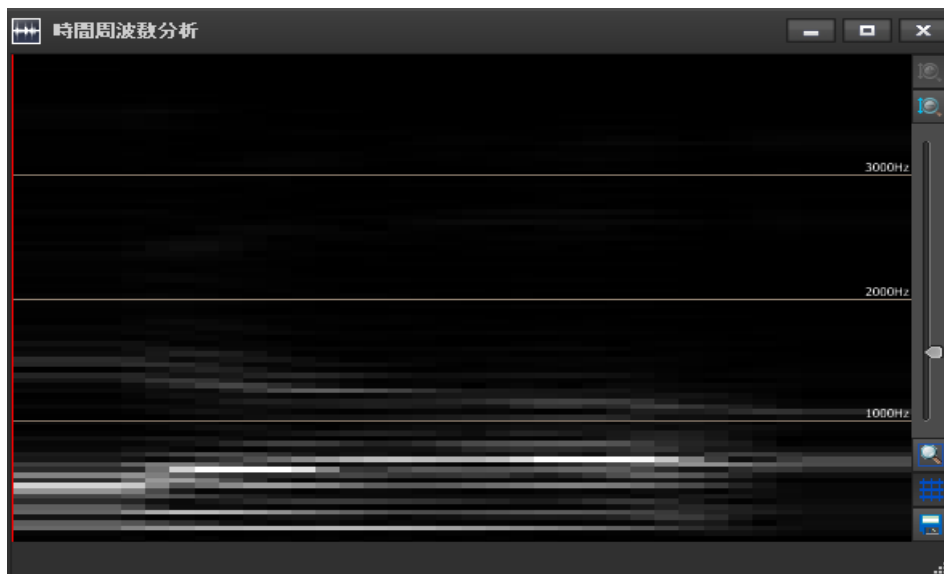


図 2.1-1 自然音声のソナグラム

縦軸が周波数、横軸が時間、濃淡が振幅（音圧）を表す。

## 2.2 音節について

文の構成単位は単語であり、単語は音節 (syllable) で構成されている。音節は、周波数変調 (frequency-modulation: FM) 成分である子音 (consonant) と、定常周波数 (constant frequency: CF) 成分の母音 (vowel) からなる (図 2.2-1)。

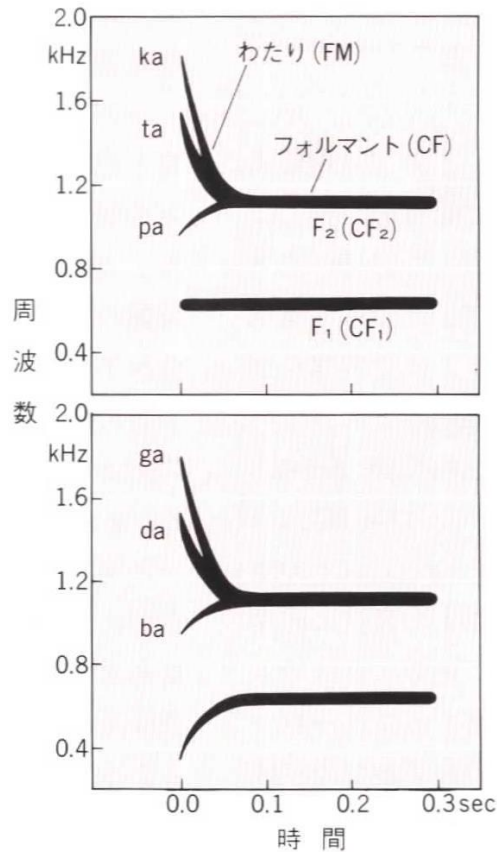


図 2.2-1 音節のソナグラム[1]

子音は周波数が変化し、母音は周波数が一定

子音は約 50ms 程度で周波数が急激に変化する。一方、母音は第一フォルマント、第二フォルマント、第三フォルマントのようにいくつかの倍音を持っており、母音によってフォルマント部分の周波数が異なる (図 2.2-2)。母音の認識には、これら 3 つのフォルマントが特に重要である。



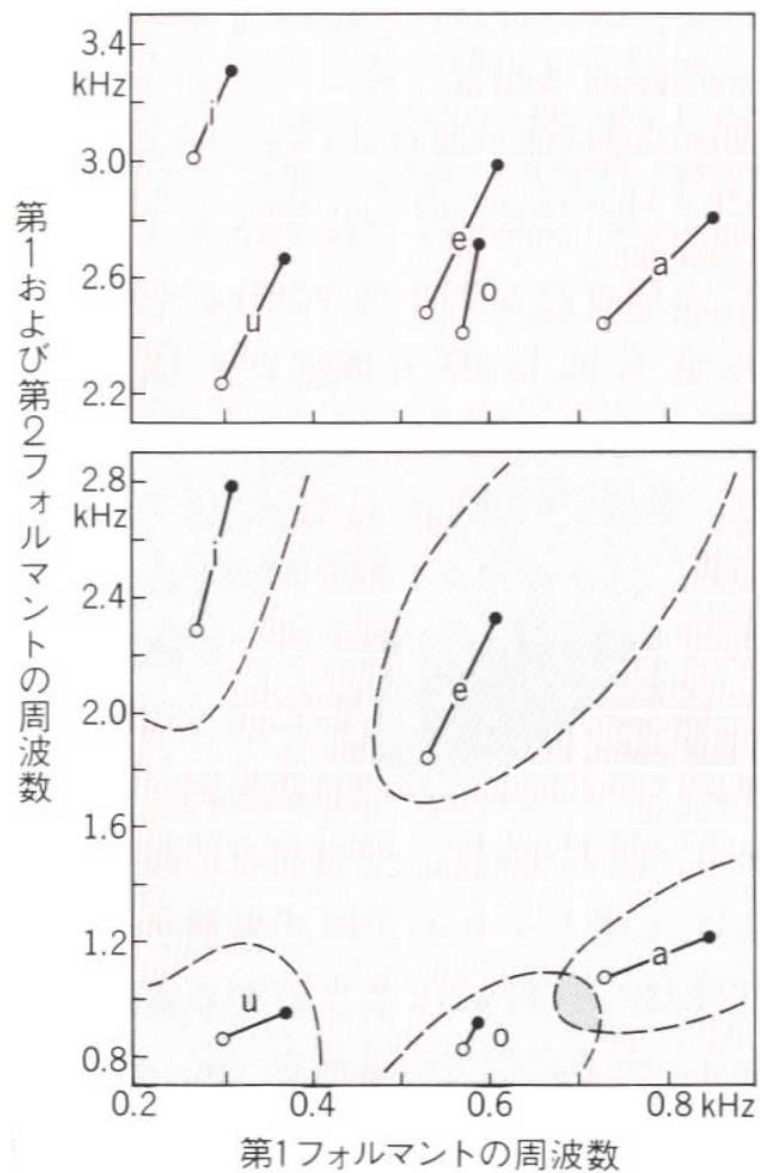


図 2.2-2 母音の第一、第二フォルマンツ [1]

○…男性のフォルマンツ周波数の平均値

●…女性のフォルマンツ周波数の平均値

## 2.3 音の受容器

音の受容器である耳は、外耳、中耳、及び内耳からなる。外耳は収音器のような役割を果たし、音を集めて中耳に送り、鼓膜を振動させる。中耳ではツチ骨、キヌタ骨、アブミ骨と呼ばれる3つの耳小骨が連鎖していて、この連鎖によって鼓膜を振動させる音のエネルギーが内耳へ送られる。内耳は非常に硬い骨に埋まっており、蝸牛、卵形嚢、球形嚢、三半規管からなる。蝸牛は前庭階、鼓室階、中心階の三室があり、前二者は外リンパ液で、後者は内リンパ液で満たされている。

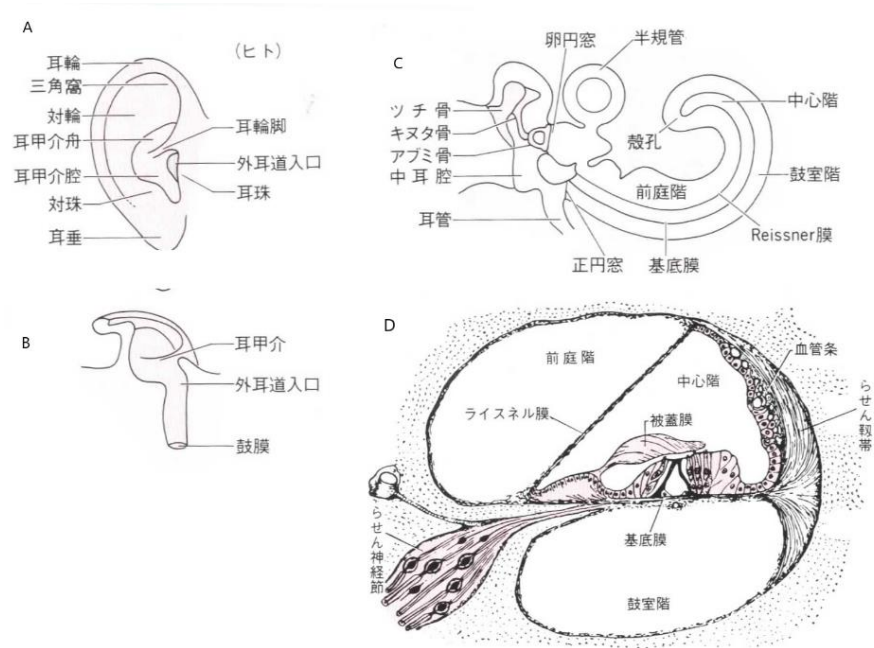


図 2.3-1 ヒトの耳[1]

A：外耳の表面図

B：外耳の断面図

C：中耳および内耳の模式図

D：蝸牛管の断面図

鼓膜の振動は、前庭階と鼓室階の外リンパ液に伝えられ、基底膜を振動させる。基底膜の上にはコルチ器官があり、コルチ器官内には内有毛細胞が一行に、外有毛細胞が三列に並んでいる。これらの有毛細胞は上部と底部がそれぞれ網状板と支持細胞とによって基底膜上に支えられている。有毛細胞の上面に生えている100~120本の不動毛はコルチ器官を覆う被蓋膜に向かって伸びているが、外有毛細胞の毛の中でも最も長いもの(30~40本)だけが被蓋膜と接している(図 2.3-2)。

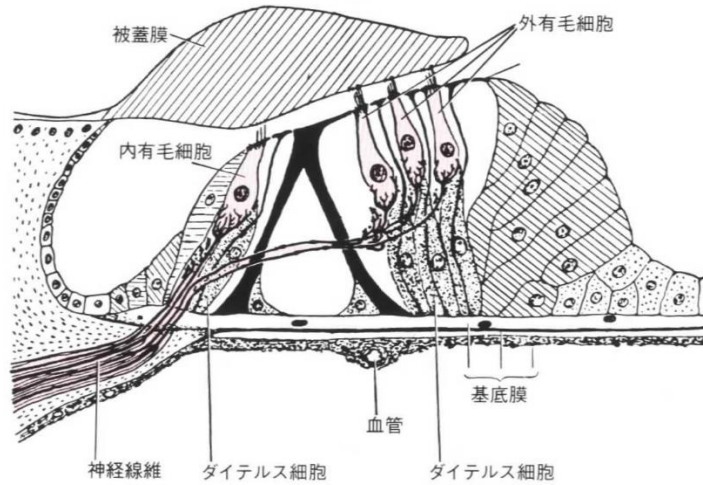


図 2.3-2 ヒトのコルチ器官の断面図[1]

内耳に伝達された音のエネルギーによって基底膜が被蓋膜へ向かって動かされると、被蓋膜に接している外有毛細胞は基粒のある方向へ傾く（図 2.3-3）。一方、被蓋膜に接していない不動毛は被蓋膜と有毛細胞上面との間にある内リンパ液が動くとき傾けられる。外有毛細胞は基底膜の動く距離と速度の両方に、内有毛細胞は速度だけに感度が良い。このように音のエネルギーが有毛細胞に伝えられる。

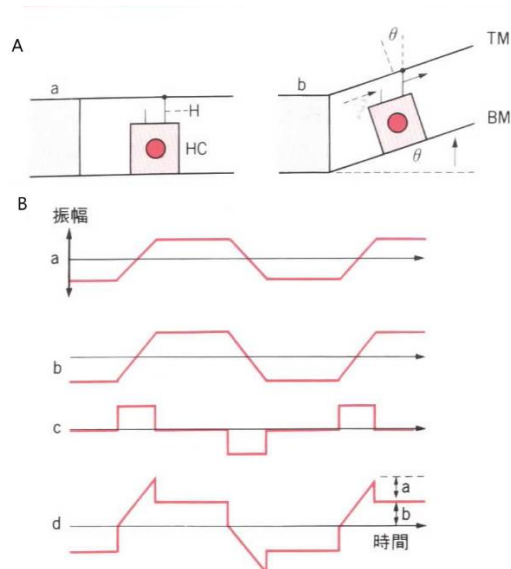


図 2.3-3 不動毛の反応[1]

A : (a) 静止状態 (b) 基底膜が動いた状態

B : (a) 刺激の波形 (b) 振幅に比例した不動毛の反応  
(c) 速度に比例した不動毛の反応 (d) 二種類の毛の反応の合計

音刺激により基底膜が被蓋膜へ向かって動くとき、イオンの透過性が増し、有毛細胞に脱分極が、逆方向へ動くとき、イオンの透過性が減り、過分極が起こる。イオンの透過性が増すと受容器電流が不動毛の先端から有毛細胞内に流れ込み、細胞膜を横切って流れ出て、細胞の脱分極を起こす。この脱分極により、有毛細胞のシナプス前終末から聴ニューロンへと興奮性の伝達物質が送られる。

## 2.4 聴覚経路

図 2.4-1 にネコの聴覚伝達路を示す。蝸牛を出た一次神経は蝸牛神経を構成して脳の中へ入り、蝸牛神経核に終わる。蝸牛神経核の神経細胞の軸索は、一部は動側の上オリーブ核群に至り、また一部は腹側聴条、中位聴条および背側聴条を通して対側の上オリーブ核群に至る。上オリーブ核群はこのように左右の耳からの入力が始めて収束する場所で、音の方向知覚に重要な役割を果たしていると言われている。上オリーブ核群の神経細胞の軸索は外側毛帯を形成して橋を上向き、中脳の下丘に終わる。蝸牛神経核の軸索の一部は上オリーブ核群で中継されることなく、直接対側の下丘に至る。下丘は脳のあまり発達していない下等な生物では聴覚系の最終点である。ネコなどの高等動物においても、下丘までの聴覚系において一部の情報処理は完成していて、大脳聴覚野を破壊した動物も、音の開始、強度の変化、周波数の変化などの音の簡単なパラメータに関する弁別の再学習をすることができると報告されている。下丘の神経細胞の軸索は下丘腕を形成して上行し、視床の内側膝状体で中継された後、大脳側頭葉にある聴覚野に至る。

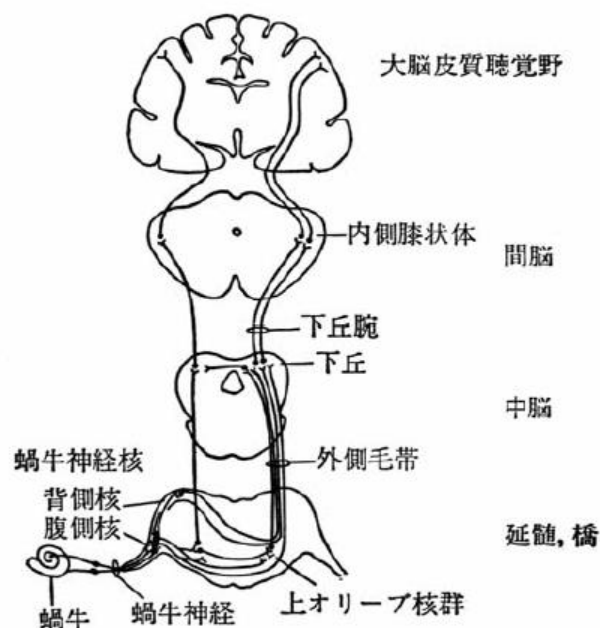


図 2.4-1 ネコの聴覚中枢系路[2]

### 3. 研究背景

#### 3.1 一次聴覚野

Yamaguchi ら[3] は図 3.1-1 に示すように guinea pig に麻酔を施し、聴覚皮質をむき出しにして電位感受性色素を用いて一次聴覚野の光学測定を行った。図 3.1-2 は測定を行った左脳の聴覚皮質の模式図である。聴覚皮質は  $12 \times 12$  の光ダイオードで聴覚皮質が発した光の信号を検知する。信号はアナログ信号をデジタル信号に変換する回路を通してデジタル信号に変換し、コンピュータで聴覚皮質の輝度解析を行った。

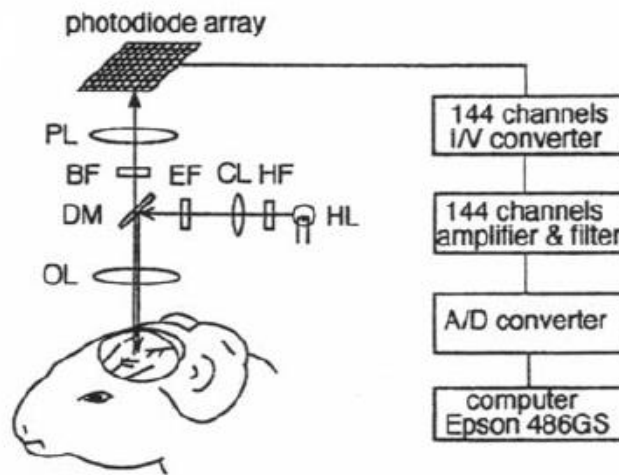


図 3.1-1 guinea pig を用いた一次聴覚野の光学測定の様式図[3]

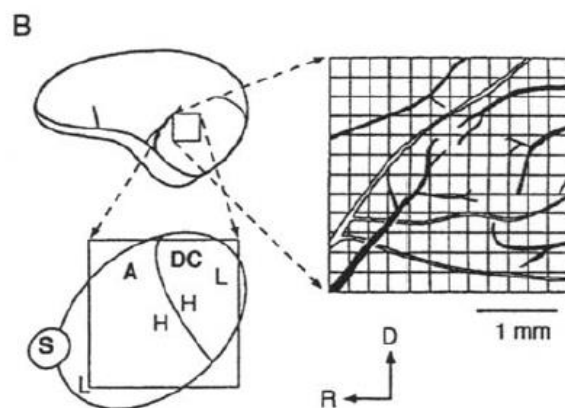


図 3.1-2 guinea pig の左脳の聴覚皮質の様式図[3]

A : anterior field, DC : dorsocaudal field, H : high frequency  
L : low frequency, D : dorsal, R : rostral

興奮性の応答は、A 領域と DC 領域で興奮性の帯域を形成する。興奮性の帯域の位置は周波数に沿って A と DC 領域を移動した。トーンバースト音を刺激として用いたときは、一時聴覚野では活性化している場所が周波数帯に沿って時間とともに動く。この結果から、時間とともに変化する音情報は、一時聴覚野においても時空間的に活性の伝播によって表現されていることがわかる。

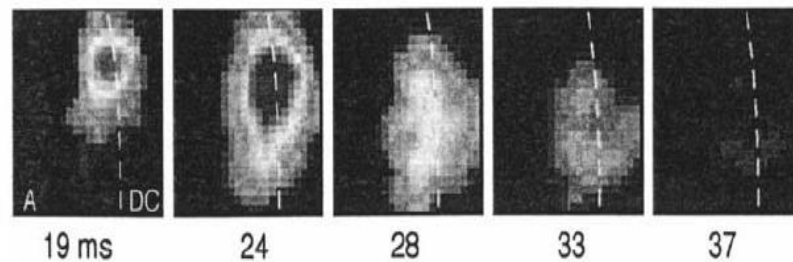


図 3.1-3 光学測定による興奮性の帯域[3]

16 kHz のトーンバースト音を刺激として用いた。下の数字は刺激を入れた時間を開始時刻とした経過時間である。破線は A 領域と DC 領域の境界線。

また、16 kHz だけでなく様々な周波数の刺激で同様の実験を行った結果をまとめた図が図 3.1-4 である。この結果より、それぞれの周波数に対して特異的に応答する領域が存在していることがわかる。これは、一次聴覚野においても、周波数の部位的表現を行っていて、その周波数帯に応じた場所で時間的に活動が伝播していることを示している。

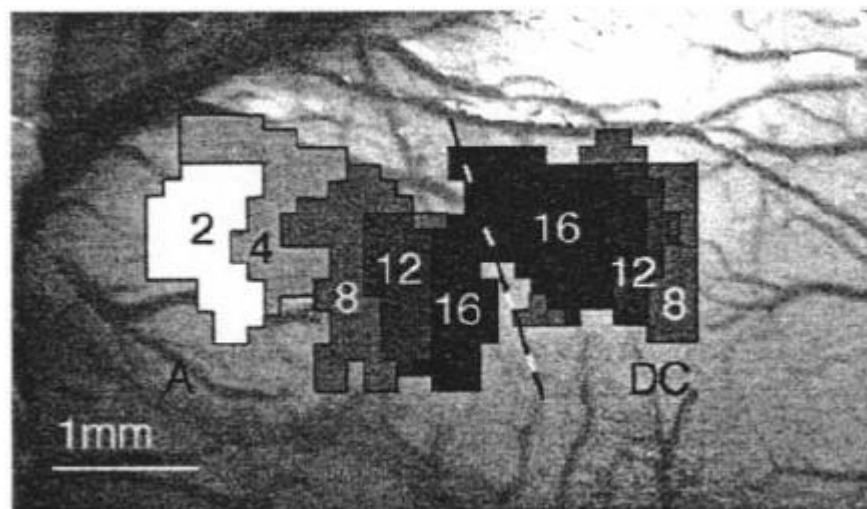


図 3.1-4 一次聴覚野における周波数の部位的表現[3]

数字は周波数 (kHz) を表し、それぞれ対応した部位が活性を示した。

聴覚皮質において、過渡的な興奮性の応答の後に抑制性の応答が続いて起こる。図 3.1-5 に示すように、抑制性の領域は興奮性の周りに位置している。この抑制性は側抑制として働き、興奮したニューロンの周りを抑制する働きがある。一次聴覚野における抑制は興奮性の領域の反応が伝播した後にそれに続くように抑制性の反応も伝播していく。

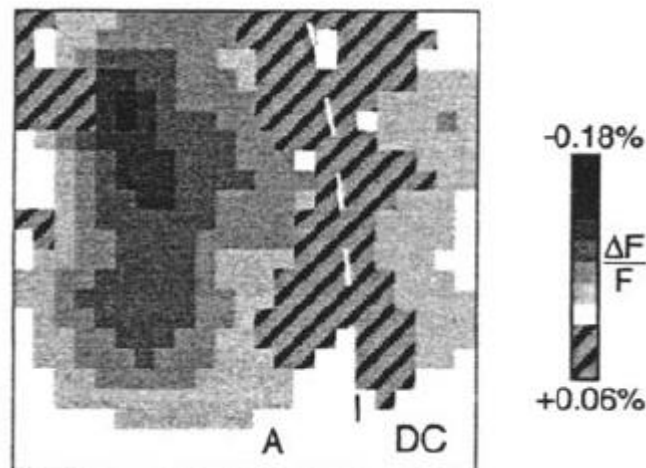


図 3.1-5 興奮性、抑制性の応答位置[3]  
斜線の領域が抑制性の応答を示している。

また、Yamaguchi ら[3]は、時間的に変化する音に対して一次聴覚野がどのような応答を示すかを研究するため、周波数が変化する FM 音を用いて実験を行った。FM 音は 4 kHz から 16 kHz まで変化する上昇音 FMa と、それとは逆に 16 kHz から 4 kHz まで変化する下降音 FMd を刺激音として用いた。FMa を刺激として用いた一次聴覚野での応答は、最初は 4 kHz の周波数帯で応答を示し、周波数が上昇するとともに反応する周波数帯も変化した。FMd の場合も同様に、最初は 16 kHz の周波数帯が応答し、最終的には 4 kHz の周波数帯まで下降音に沿って応答部位が変化した。

次に、Yamaguchi ら[3] は連続的な周波数変化をする FMa や FMd と違い、音の周波数を段階的に変化させた。刺激音 SFa は 4-8-16 kHz を 100 ms 用いた。SFa の応答は、それぞれの周波数は同じ時間遅れで、4、8、16 kHz の周波数帯で現れた。このパターンは、純音を刺激として用いたときと似た応答を示した。

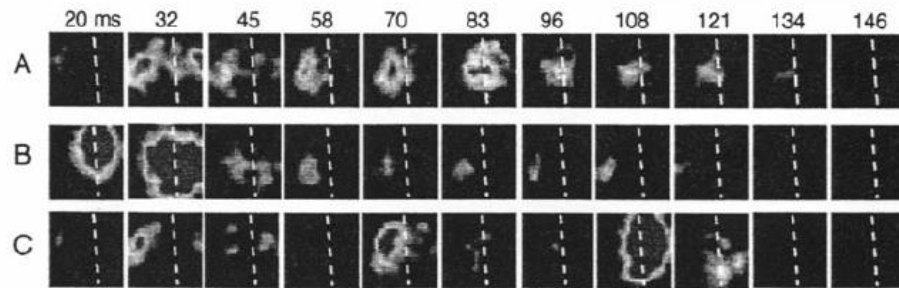


図 3.1-6 FM 音を刺激として用いたときの一時聴覚野の応答[3]  
A は FMa、B は FMd、C は SFa を刺激として用いたときの応答。  
時間は入力からの経過時間である。

純音の一時聴覚野の応答は、周波数帯に沿って伝播し、FM 音は周波数帯を斜めに横切って伝播する、という違いがあることがわかった。

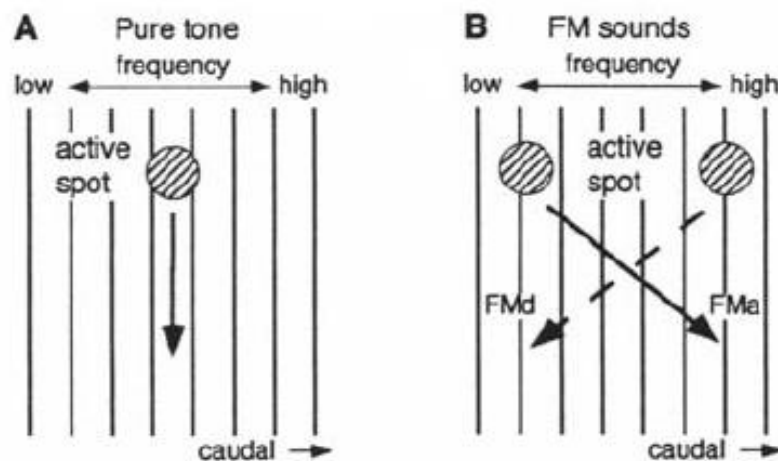


図 3.1-7 純音と FM 音の活性の伝播[3]  
A は純音を刺激として用いたときの伝播の様子を、B は FM 音を刺激として  
用いたときの伝播の様子を模式図で表したもの。

山口、谷口は、純音刺激が聴覚皮質の周波数帯に垂直な方向に伝播するという性質から、周波数軸（トノトピー軸）と、それに対して垂直な伝播軸の 2 次元平面で構成される神経回路モデルを提案した[4]。このネットワークは興奮性のニューロンと抑制性のニューロンの 2 つからなるユニットを 2 次元的に配列したネットワークである。





図 3.1-8 興奮性ニューロンと抑制性ニューロンからなるユニット[4]  
 $x$  は興奮性のニューロン、 $y$  は抑制性のニューロン。

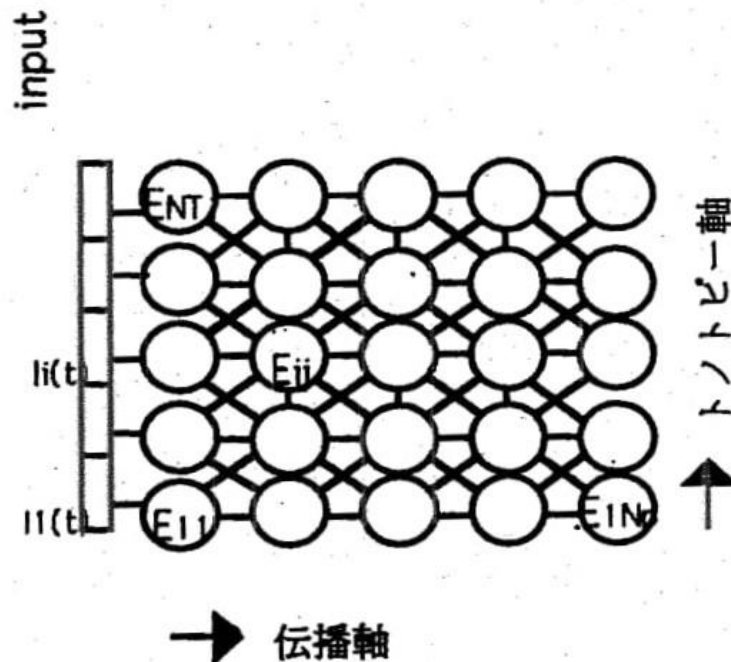


図 3.1-9 一時聴覚野の 2 次元ネットワークモデルの概要図[4]  
 周波数軸方向は  $\pm 1$  ユニット、伝播軸方向は  $+1$  ユニットと結合している。

この神経回路に純音 (CF 音)、FM 音を入力したとき、Yamaguchi らが行った実験で示されたような一時聴覚野の性質を再現した。直接入力を受けた左端のユニットは右に並ぶ複数のユニットに活動を伝えようとする。活性化されたユニットはさらにその右へ伝達し、次々と活性が右へずれていく。つまり、周波数軸の上ではその瞬間に入力のあるものが選ばれて活性化し、伝播軸上ではそのユニットの隣が活性化することになる。このようにしてネットワーク上を波動的に伝播していく。

CF 音を入力した場合、等周波数帯における一方向の伝播が繰り返し起こる (図 3.1-10)。活動の大きさの経時変化を調べると、1 回目の伝播活動に比べて 2, 3 回目の伝播活動は減少していく傾向がみられる。これは、1 回目以降で起こる抑制系の活性化が原因である。

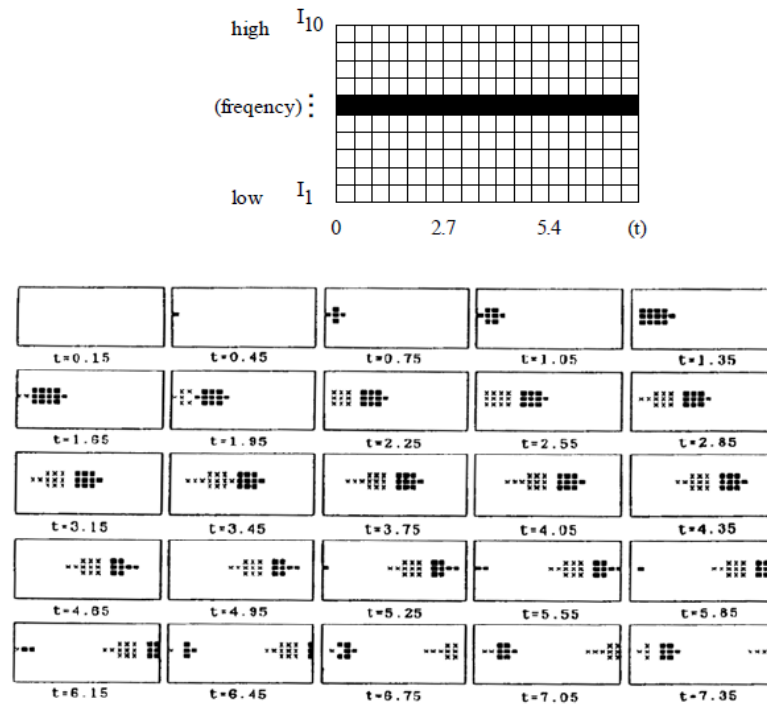


図 3.1-10 CF 入力時の活動パターン[4]

■：活性化しているユニット    ×：過分極しているユニット  
 ブランク：静止状態に近いユニット

一定の変調のかかる FM 音を入力した場合、低→高または高→低を入力すると実験でみられるように回路中を斜めに横断する活動伝播が得られる。また、低→高→低→高のような FM 音の連続入力の場合、一様な周波数変調の時期はほぼ一定の活動領域が移動するのに対し、変調方向が逆転する時点では活動領域の縮小がみられる (図 3.1-11)。

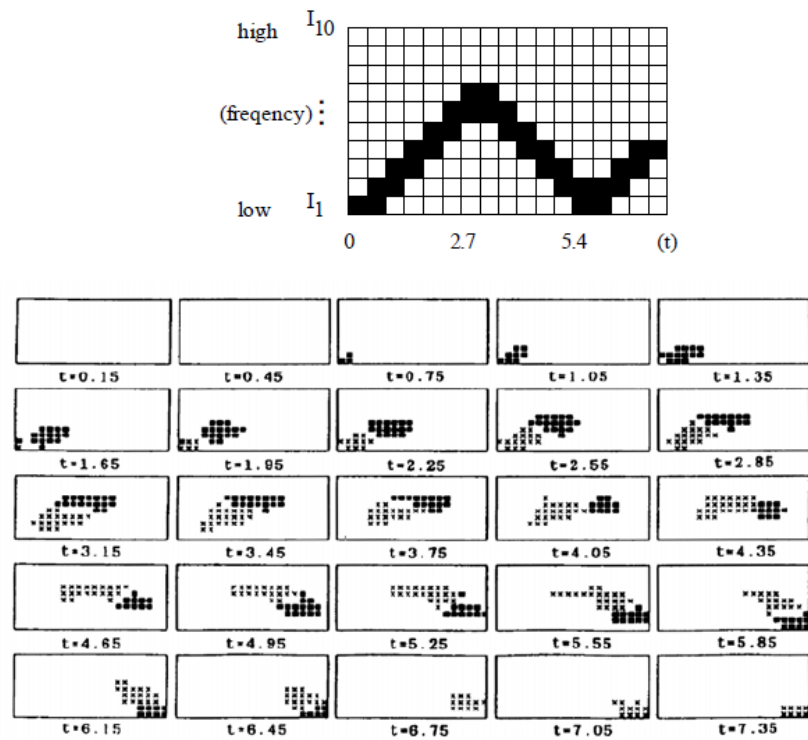


図 3.1-11 FM 音入力時の活動パターン[4]

■：活性化しているユニット    ×：過分極しているユニット  
 ブランク：静止状態に近いユニット

### 3.2 音情報のカテゴリー化

Edward F Chang ら[5]はヒトが音声を聞き取る際に音をカテゴリー化している領域を発見した。この領域は言語野の前側頭葉に存在する。

実験では、ヒトに対して/ba/、/da/、/ga/といった、ソナグラム的にも近い14の音を区別するタスクを課し、そのときの脳波を測定したものである。図 3.2-1 は実験で用いた音刺激の例であり、図 3.2-2 はタスクで使用された音刺激がそれぞれ/ba/、/da/、/ga/のどの音に分類されるかを示した図である。

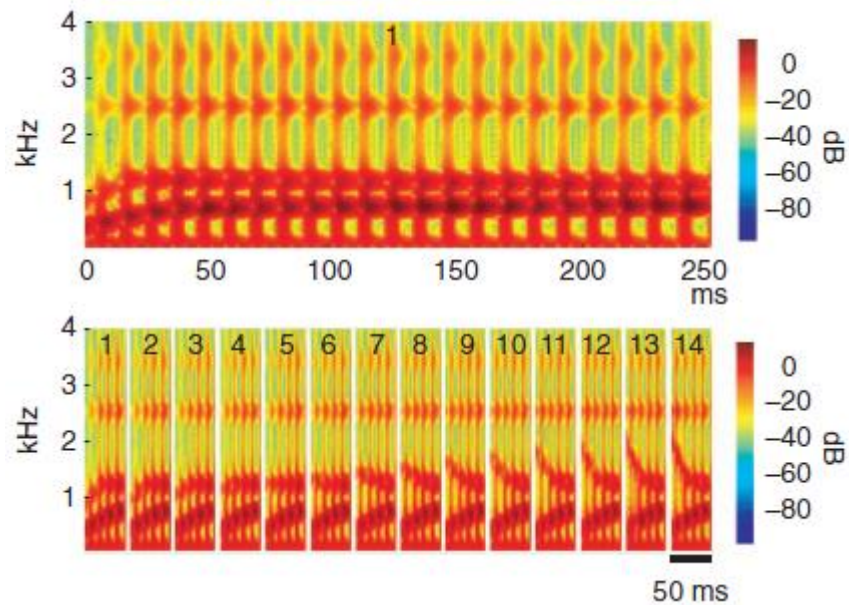


図 3.2-1 用いた刺激音のソナグラム例[5]

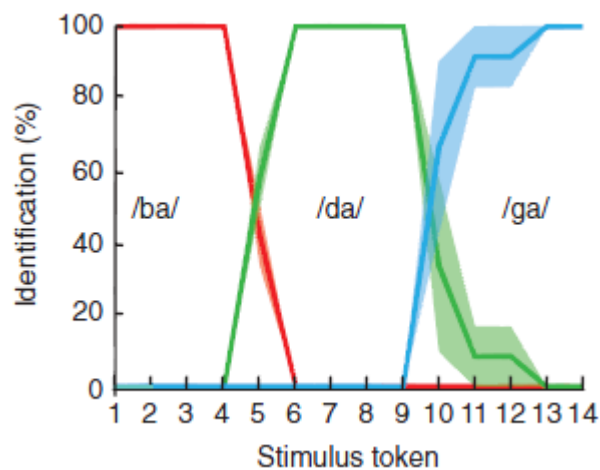


図 3.2-2 刺激音の分類[5]

14 の刺激音が/ba/、/da/、/ga/のどれに分類するかを示した図。  
例えば刺激音 3 なら/ba/、刺激音 5 なら/ba/と/dɑ/の中間の音である。

これらの刺激音を用いて前側頭葉の脳波を計測した。3種類それぞれの音を聞いたとき、そのピークはそれぞれ異なった領域が活性しているという結果が得られた。

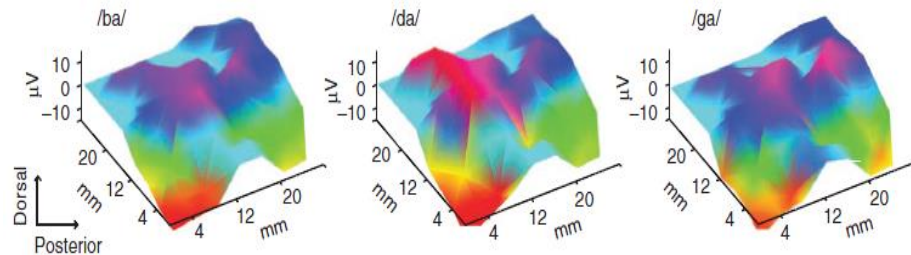


図 3.2-3 刺激音を与えたときの前側頭葉の応答[5]

前側頭葉の応答を三次元的に表した。xy 平面は前側頭葉の領域を、z 軸方向はその活性度をそれぞれ表している。

また、/ba/、/da/、/ga/それぞれの刺激音に対する反応だけでなく、14種類の刺激音のうち2つを比較し、どれに分類されるかというタスクを課し、そのときの前側頭葉での応答を観測した。まず、前側頭葉が刺激音を入力されてからどのくらいの時間でカテゴリー化を行っているのかを調べた。図 3.2-4 に示す通り、カテゴリー化を行うニューロンは刺激を入れて即座に反応はせず、約 110 ms 経過したときに最も反応している。

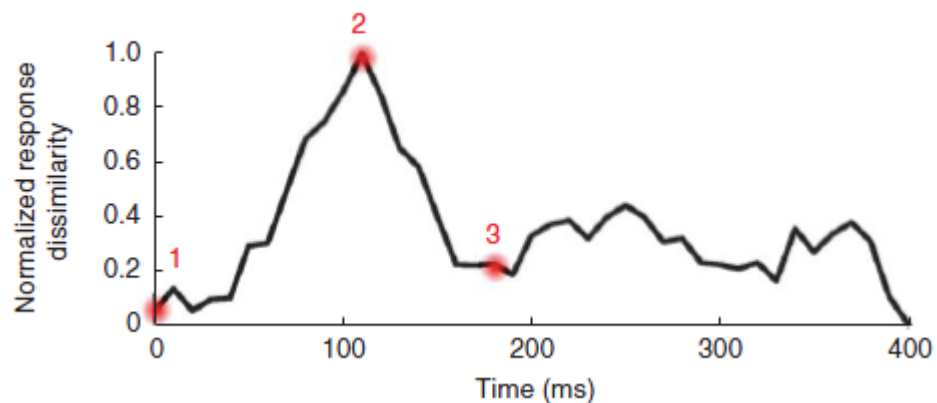


図 3.2-4 前側頭葉の活性の経時変化[5]

また、これら3つの音のカテゴリー化では、前側頭葉内でも様々な場所で応答している。図 3.2-5、図 3.2-6 に示すように、カテゴリー化を行っている領域が重なっている領域もあり、スパース表現をしていることが示唆されている。重なっている領域は被験者により個人差はあるものの、5%以内でそれほど多くはない。

が、同じニューロン群を用いて複雑にカテゴリー化をコードしている可能性がある。

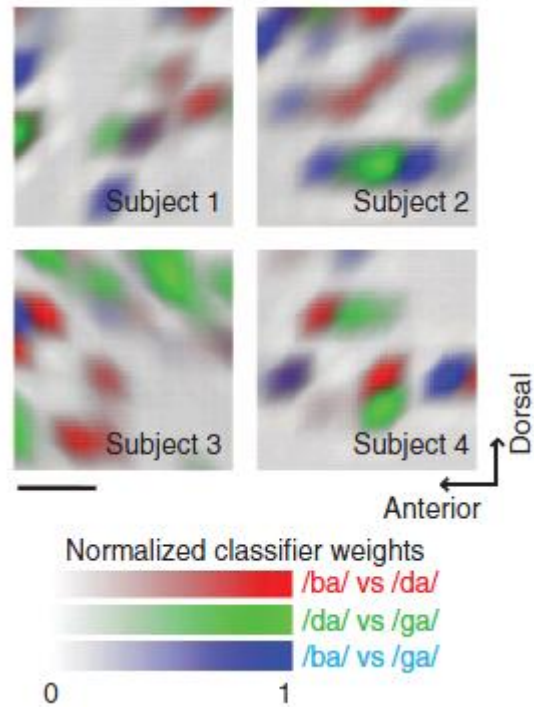


図 3.2-5 前側頭葉におけるカテゴリー化[5]

4 人の被験者に対して /ba/、/da/、/ga/ の分類のタスクを行った際のそれぞれの前側頭葉の応答。

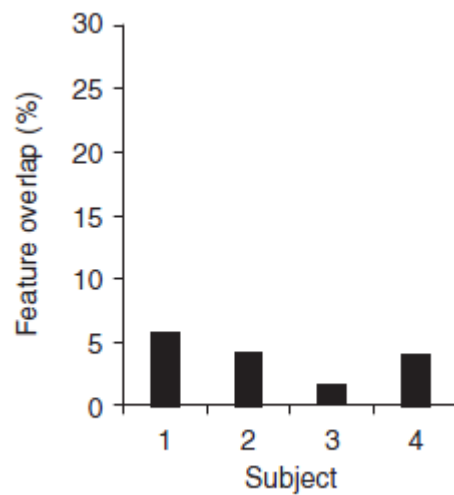


図 3.2-6 前側頭葉における応答の重複[5]

4 人の被験者に対して /ba/、/da/、/ga/ の分類タスクを行った際に重複した反応を示した割合。

### 3.3 言語処理と脳内リズム

神経振動（リズム）は脳内のいたるところにあり、いくつかの方法で認知に貢献するといわれている。例えば、情報を分離し、スパイクタイミングを整理することが挙げられる。聴覚認識では、音声は連続的に情報が入力されるが、そのような音響信号を適切な時間粒度の単位にセグメント化する必要があるといわれている。

Pefkou ら[6]は、被験者に対して圧縮されたフランス語を聞かせ、音声の理解度と脳内リズムの関係性について調べた。図 3.3-1 に、実際に行われた実験を示す。

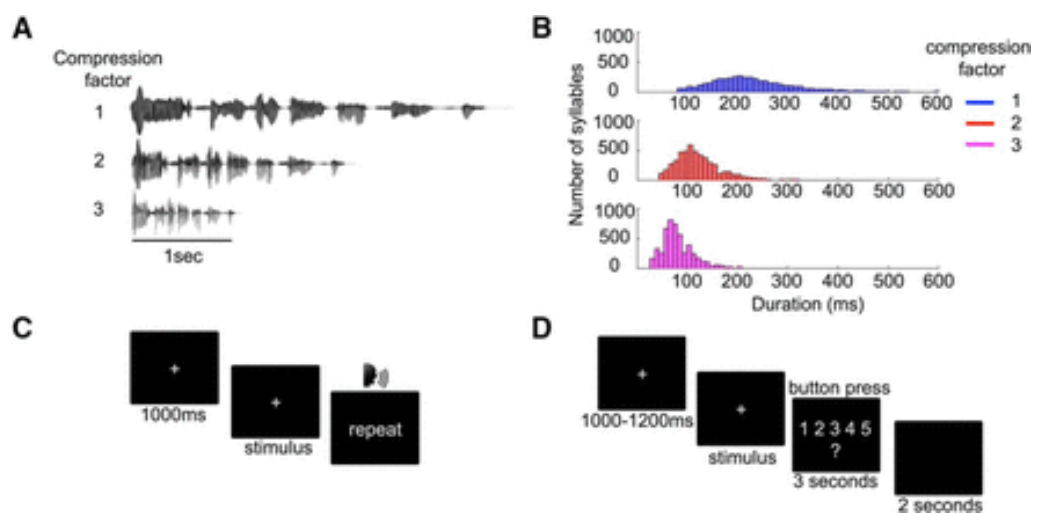


図 3.3-1 音声波形と行動タスク[6]

A：フランス語の波形。

1 が元々の波形、2、3 はそれぞれ 2 倍、3 倍に圧縮されたもの。

B：音節の持続時間。それぞれの音節の長さの分布を表す。

C：行動タスク。聞いた音声を被験者が繰り返す。

D：脳波記録中の行動タスク。聞いた音声の理解度を 1 ～ 5 で評価する。



圧縮されていない音声、2 倍圧縮の音声については理解度に大きな差はないが、3 倍圧縮になると著しく理解度が下がる（図 3.3-2）。

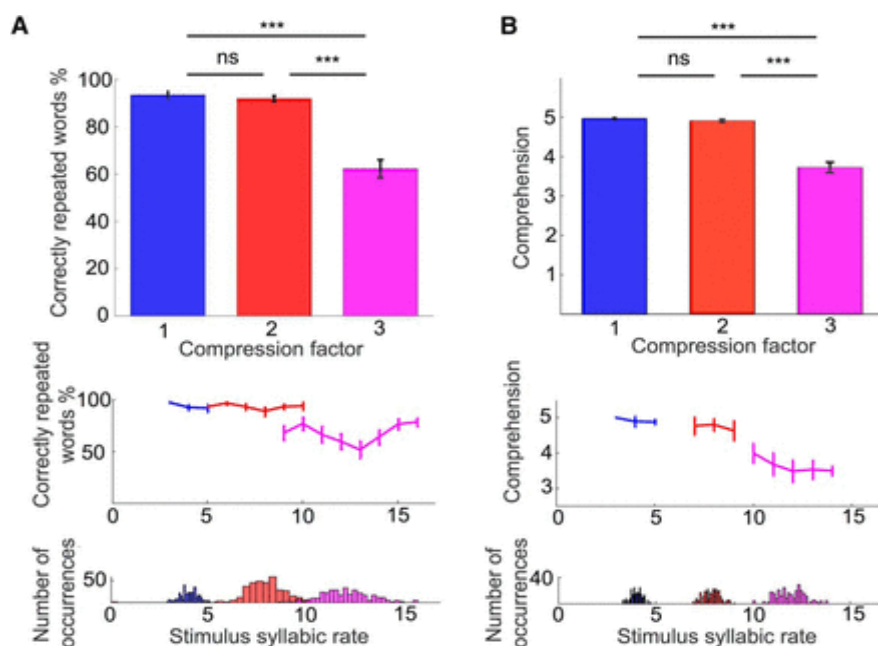


図 3.3-2 行動タスクの結果[6]

A：行動タスクの結果 B：脳波記録中の行動タスクの結果  
上段・中段：正答率及び理解度 下段：音節の発生率

脳波記録中の行動タスクについては、それぞれどの周波数の振動が、どのようなときに活性化されるかについても調べられている（図 3.3-3）。theta 振動（4 Hz）のパワーは音節速度に応じて増加したが、理解度との関連は見られなかった。逆に beta 振動（14-21 Hz）のパワーは音節速度が速くなると低くなり、理解度が高いほど beta パワーも高い。

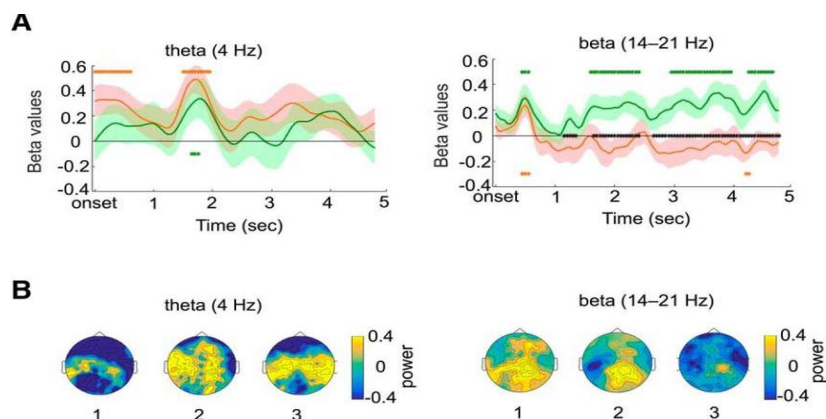


図 3.3-3 音節速度と理解の効果の時間経過[6]

A:回帰時間経過 B:頭皮トポグラフィー



また、Giraud ら[7]は、安静時とフランス語の音声を聞いたときの、ヒトの聴覚皮質活動の時間周波数表現について調べた。図 3.3-4 に、Giraud らが行った実験結果を示す。

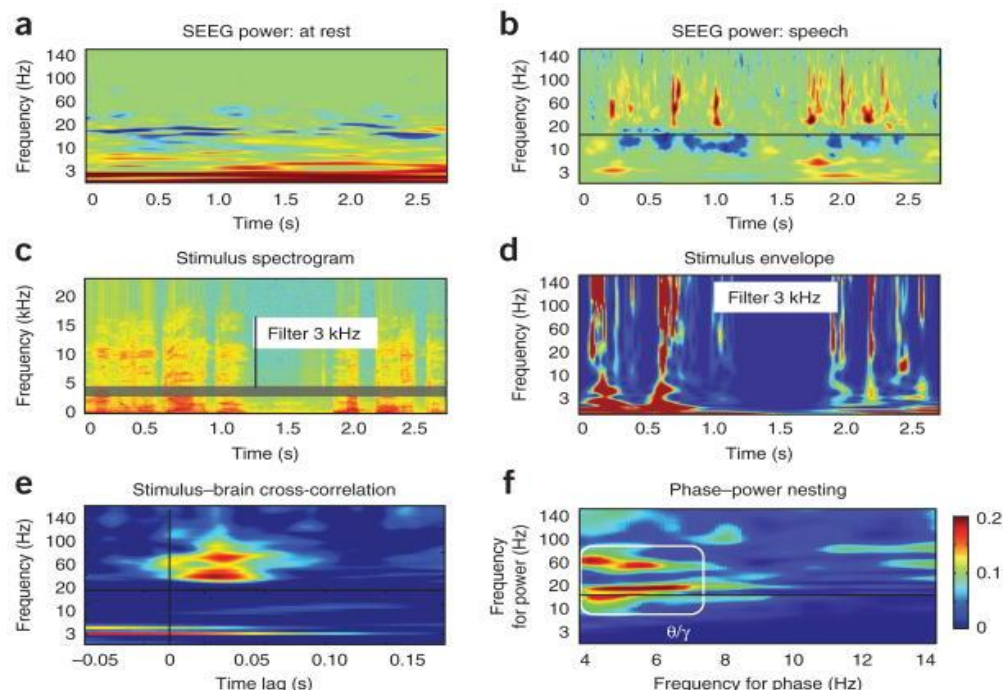


図 3.3-4 実験結果[7]

- (a) 安静時の皮質活動の時間周波数表現
- (b) フランス語の音声文「Le nouveau garde la porte」への  
応答における皮質活動の時間周波数表現
- (c) 刺激のスペクトログラム
- (d) c の 3 kHz を中心とする帯域から抽出した変調スペクトル
- (e) 刺激と皮質活動との相互相関
- (f) 刺激と脳の間周波数固有のロックを反映する試行間位相一貫性の指標

安静時の皮質活動は、主に 1-25 Hz の周波数帯域で持続的な振動活動をしている。聴覚皮質が音声によって刺激されると、活動は一時的に変動する。このとき、刺激と皮質活動との相互相関を調べると、図 3.3-4(e)のように、 $\theta$  及び低  $\gamma$  ~ 高  $\gamma$  で高い。

Giraud らは、これらの結果から、聴覚皮質の神経振動が聴覚情報の重要な要素を構成し、連続的な音情報を解読する最初のステップに関与しているという仮説を立てている。

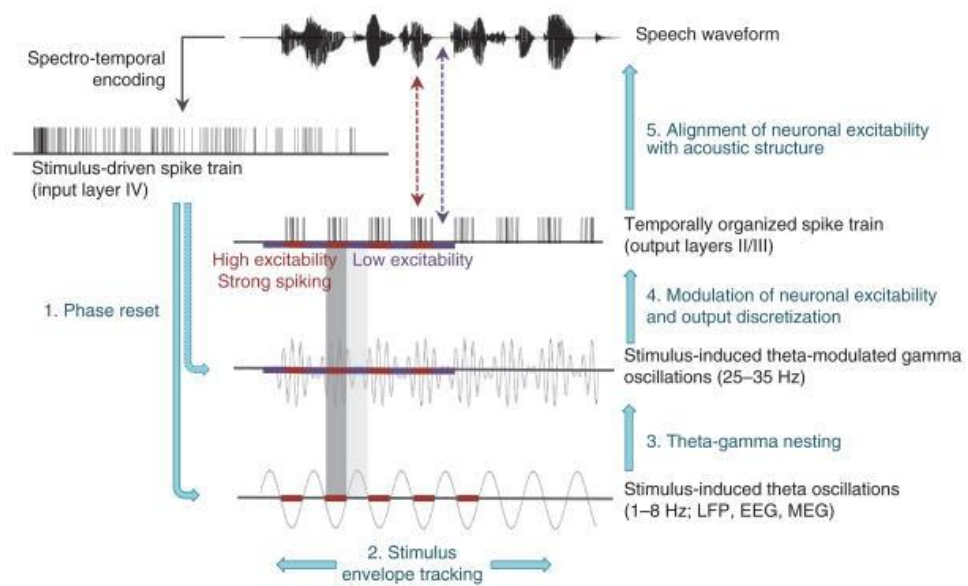


図 3.3-5 音声認識における振動ベースの初期操作の理論[7]  
あるスピーチが刺激として入力されたとき、連続で発火が生じるが、  
振動により発火を離散化し、情報を整理することができる。

## 4. 本研究の問題

### 4.1 本研究の問題

前章で述べたように、聴覚情報は脳において神経活動の時空間パターンとして表現されること、音ごとにカテゴリー化している領域があること、そして脳内リズムが音声認識に関わっている可能性があることなどが明らかになってきている。しかし、これらの知見がどのようにつながり、聴覚野で音情報が処理されているのかについてはよくわかっていない。

そこで、本研究では以下の問題について取り扱う。

1. 一次聴覚野における神経活動の時空間パターンから、どのように音情報を得ているのか？
2. そのメカニズムにおいて、脳内リズムがどのような関与をしているのか？

これらの問題を解決するため、ここでは音節情報の処理について考える。

### 4.2 アプローチ

鼓膜から聴覚経路を上行してきた音情報を一次聴覚野にてニューロン活動の時空間パターンに変換する。その様々なパターンから Kohonen map を用いて特徴を抽出する。その上層に音節情報をまとめ上げるニューロン群を作成し、音節情報のまとめ上げを行う。これらの機能を3つの層によって階層的に処理することによって音声認識を行うニューラルネットワークモデルを作成し、音声認識の神経機構を解明する。

また、各層における音情報表現に脳内リズムが関わっていると考えた。ガンマ波(30~120Hz)は、最初の時空間パターンで音の基本情報をコードすることに関係している。また、ベータ波(10~20Hz)は、特徴抽出の段階で、子音、母音に関わる情報表現を取り出すことに寄与している。さらに、シータ波(~10Hz)は、子音と母音をまとめたシラブル表現に関わっている。これらのリズムの効果を上記モデルに加えて、音の各表現を作り出すときに各リズムがどのように役立っているかを調べる。

## 5. モデル

### 5.1 モデル概要

本研究で作成したモデルは、Fujita et al. [8] が提案したモデルに基づいている。図 5.1-1 に示すように、3 層構造になっている。すべての層が 2 次元のニューラルネットワークからなり、各層はフィードフォワード結合によってつながっている。

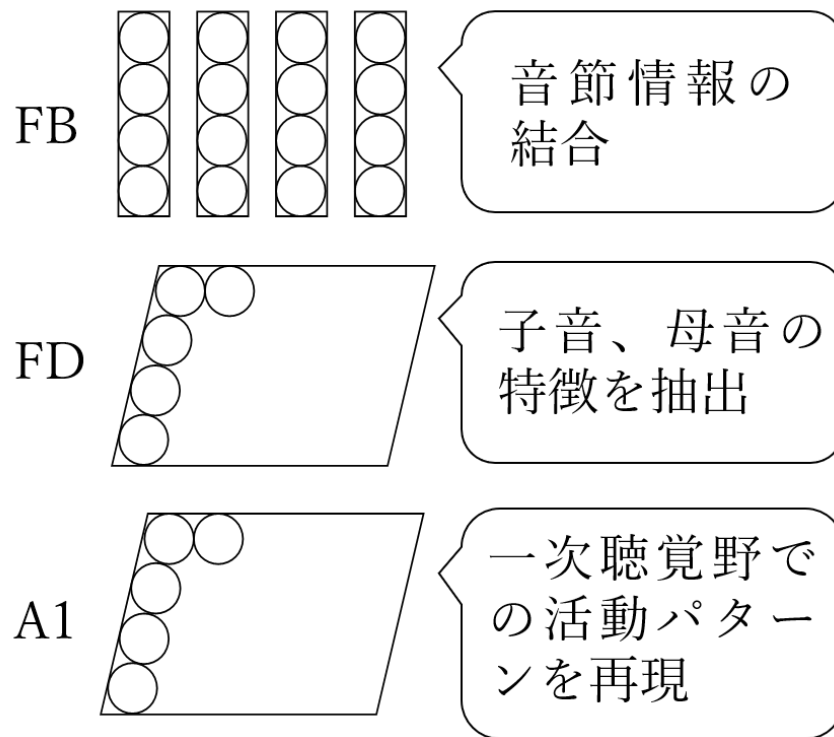


図 5.1-1 モデルの概要

図 5.1-1 に音情報処理の各層の機能的役割を示す。A1 層は一次聴覚野を表現していて、音信号を神経活動の時空間パターンに変換する。FD 層 (feature-detective layer) では、A1 層の活動パターンから Kohonen map を用いて子音や母音などの特徴を抽出する。FB 層 (feature-binding layer) では、FD 層で得た子音や母音の特徴を結合し、音節 (シラブル) としての情報にまとめ上げる。

## 5.2 一次聴覚野 (A1 層)

A1 では、Yamaguchi ら [3, 4] の実験結果から音声信号を時空間的な活動パターンに変換するという役割が明らかになっている。山口、谷口 [3, 4] は A1 をモデル化し、音信号からニューロン活動の時空間パターンの形成を示している。本研究では、この山口・谷口モデルを改良したものを用いる。

A1 層では、活性の伝播を表現するため、一過的自己増幅的活動をするユニットを考えた。図 5.2-1 にそのユニットの構造を示す。x ニューロンは興奮性ニューロン、y ニューロンは抑制性のニューロンである。

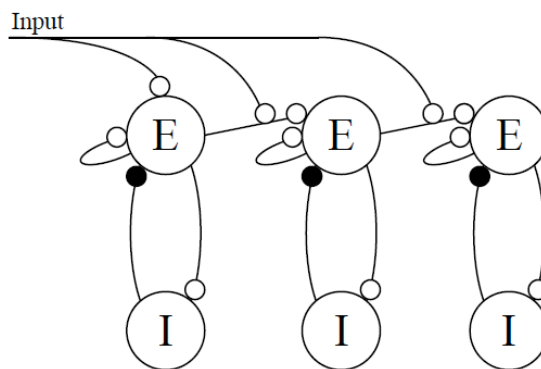


図 5.2-1 A1 層を構成するユニット列

A1 層は、このユニットを 2 次元格子状に配列することによって活性の伝播を表現する (図 5.2-2)。図 5.2-1 に示すように、周波数 (トノトピー) 軸上で  $i$  番目、伝播軸上で  $j$  番目のユニットを  $E_{ij}$  と表す。このユニット  $E_{ij}$  に対して、 $E_{k,j-1}$  ( $k=i-1, i, i+1$ ) のユニットが結合する。

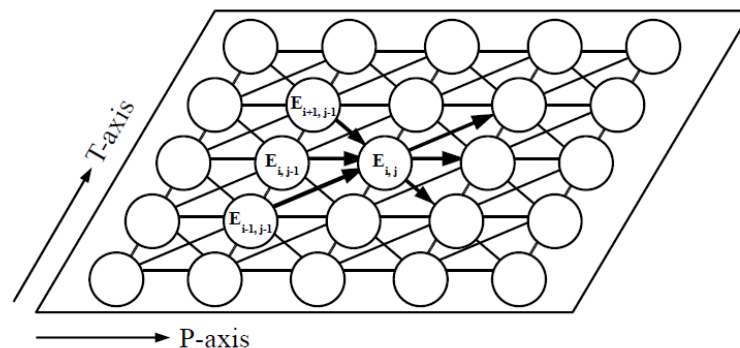


図 5.2-2 2 次元回路モデル

ニューロンの活性は近くのニューロンに伝わり、活性が伝播していく。

ユニットの方程式は、興奮性ニューロンの活性度  $x$  と、抑制性ニューロンの活性度  $y$  より与えられる。

$$\tau \dot{x}_{ij} = -x_{ij} - \alpha P(y_{ij}) + \kappa P(x_{ij}) + \Gamma + \sum_{m,n} w_{ijmn} P(x_{mn}) + \delta_j I_i - H_{ij} \quad (1)$$

$$\delta_i = 1, \text{ if } j=1, \text{ and } 0 \text{ otherwise}$$

$$\tau \dot{y}_{ij} = -y_{ij} + \beta P(x_{ij}) - \beta_0 \quad (2)$$

$$\text{with } P(z) = 0.5 \tanh\{\lambda_z(z - \mu_z)\} + 0.5, z \in \{x_{ij}, y_{ij}\} \quad (3)$$

$$w_{ijmn} = W_0(I_i + I_m + w_0) \quad (4)$$

$$H_{ij} = h_n \sum_{m,n} P(x_{mn}) \quad (5)$$

ここで、 $P$  はニューロンの出力、 $w_{ijmn}$  はユニット  $E_{mn}$  から  $E_{ij}$  への重み、 $H_{ij}$  は 2 次元回路からの集成的抑制性入力で、回路全体の活性に依存する。また、各パラメータを表 5.2-1 に示す。

表 5.2-1

$\alpha$	16.0	$\mu$	2.0
$\beta$	8.0	$h_n$	0.025
$\beta_0$	1.0	$W_0$	2.0
$\kappa$	4.0	$w_0$	0.05
$\Gamma$	1.6	$N_T$	40
$\lambda$	9.0	$N_P$	100
$\tau$	3.0		

ネットワークの大きさ： $N_T \times N_P$  ( $40 \times 100$ )

### 5.3 Feature-detective layer (FD 層)

FD 層では、A1 層の活性パターンから、Kohonen map を用いて特徴を抽出することで音のカテゴリー化を行った。

Kohonen map は、SOM (self-organizing feature map) と呼ばれ、ランダムなニューロンのシナプス結合分布から刺激の位相地図を組織化できるものである。入力層と出力層の 2 層構造であり、入力層のすべてのニューロンが出力層のすべてのニューロンに全結合している。出力層は 2 次元配列で、各ニューロン間の結合は無い。入力データが入力層に入ること、出力層に値が送られる。このとき、最も強い入力を受け取る出力層のニューロンを勝ちニューロンと呼ぶ。開始時、2 層間の結合はランダムであるが、勝ちニューロンが決まると、勝ちニューロン付近のニューロンと入力層のニューロンとの結合が大きくなる。この作業を繰り返し、学習が進むことによって、類似した入力データは出力層において近くに表示され、違いの大きな入力データは出力層において離れたところに表示されるようになる。

今回作成したモデルでは、入力層を A1 層、出力層を FD 層とした。A1 層の  $i,j$  番目のユニットの活性度を  $e_{ij}$  とする。2 つの層に存在するユニットは上下間で全結合しており、FD 層の  $k,l$  番目のニューロンに対する結合重みを  $w_{ijkl}$  とする。入力信号がセットされると結合重みによって勝ちニューロンが決定され、勝ちニューロンを中心にガウス関数に従ってその周りの重みも更新する。これにより、一次聴覚野での似たパターンは FD 層で近い場所に特徴抽出細胞ができる。

$$S_{kl} = \sum_{ij} w_{ijkl} \cdot e_{ij} \quad (6)$$

$$\Delta w = w_{max} \exp \left( -\frac{(k-k')^2 + (l-l')^2}{2\sigma^2} \right) \quad (7)$$

$$w_{kl}^{new} = w_{kl}^{old} + \Delta w \cdot e_{ij} \quad (8)$$

ここで、 $w_{max}=0.001$ 、 $w_{ijkl}$  の初期値  $=0.0\sim0.01$  の乱数、 $\sigma=5.0$ 。

今回、 $S_{kl}$  が最大となる  $kl$  番目のニューロンを勝ちニューロンとする。 $k',l'$  は更新するニューロンの位置である。

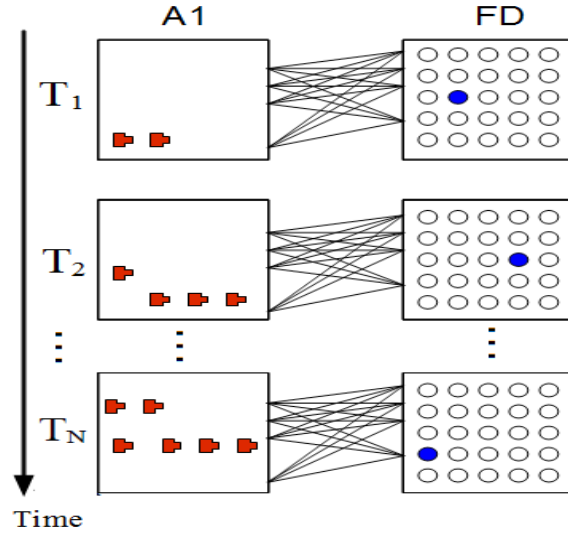


図 5.3-1 A1 層と FD 層の結合の模式図

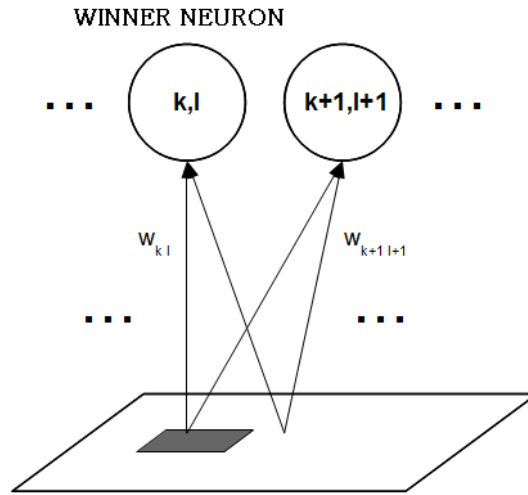


図 5.3-2 勝ちニューロンの決め方

$W_{kl} > W_{k+1,l+1}$  のとき、FD 層の  $k,l$  番目のニューロンが勝ちニューロンとなる。

FD 層の各ユニットは A1 層と同様に興奮性と抑制性ニューロンからなる。ただし、各ユニット間に結合はない。また、この層では子音や母音などの音の要素的な特徴を抽出すると考え、 $\gamma$ リズムよりも周波数の大きい $\beta$ リズム（14 - 21 Hz）を加えた。

$$\tau \dot{x}_{kl} = -x_{kl} - \alpha P(y_{kl}) + \kappa P(x_{kl}) + \Gamma + S_{kl} + A \sin(2\pi f t) \quad (9)$$

$$\tau \dot{y}_{kl} = -y_{kl} + \beta P(x_{kl}) - \beta_0 \quad (10)$$



## 5.4 Feature-binding layer (FB 層)

FB 層では、FD 層で抽出した子音や母音の特徴をまとめ、音素としての情報の結合を行う。ここでは、発火により音素を識別したか判断するため、発火型ニューロンモデルである LIF (Leaky integrate-and-fire) モデルを用いた。LIF モデルは、チャンネルの開閉や膜電位の動きなどの細かいメカニズムを考えず、入力 の時間積分によってニューロンの膜電位が変化し、ある閾値を超えると発火するモデルである。また、今回は不応期(3 ms)も考慮した。以下に、今回使用したモデルの方程式を示す。

$$I_{mn} = \sum_{kl} w_{klmn} \cdot P(x_{kl}) \quad (11)$$

$$\tau_{LIF} \frac{dV_{mn}}{dt} = -V_{mn} + I_{mn} + \sum_{m'n'} W_{m'n'mn}^{STDP} \cdot S_{m'n'} + B \sin(2\pi f' t) \quad (12)$$

$$\tau_s \frac{dS_{m'n'}}{dt} = -S_{m'n'} + \lambda_s \delta(t - t_k) \quad (13)$$

$$V = \begin{cases} V_{sp}, & (V > \theta) \\ V, & (V \leq \theta) \end{cases} \quad (14)$$

ここで、 $w_{klmn}$  は FD 層のニューロンから FB 層のニューロンへの結合、 $W_{m'n'mn}^{STDP}$  は FB 層内の結合である。FB 層では、音節情報のまとめ上げを行うため、単一音素 (シラブル) の持続時間に近い周期を持つ  $\theta$  振動を加えた。

FB 層は FD 層で抽出された子音部分 (C)、子音と母音が混在している部分 (C+V)、母音部分 (V) それぞれに反応するニューロン群で構成される。

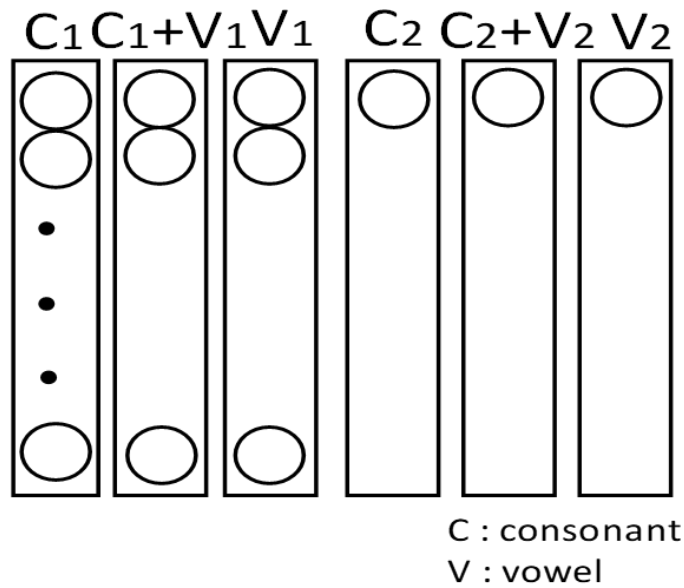


図 5.4-1 FB 層の模式図

また、同一音素内の子音、母音などのアトラクタ間の相関をつけるために、層内の結合には STDP (Spike-timing-dependent plasticity) 学習則[9]を用いた。この学習は、スパイクタイミングに依存してシナプス間の結合強度  $W$  が変化していく学習則である。ニューロン A とニューロン B があるとき、先にニューロン A が発火し、それによってニューロン B が発火したとする。このとき、 $A \rightarrow B$  間の結合強度は強まり、逆に  $B \rightarrow A$  の結合強度が弱まる (図 5.4-2)。

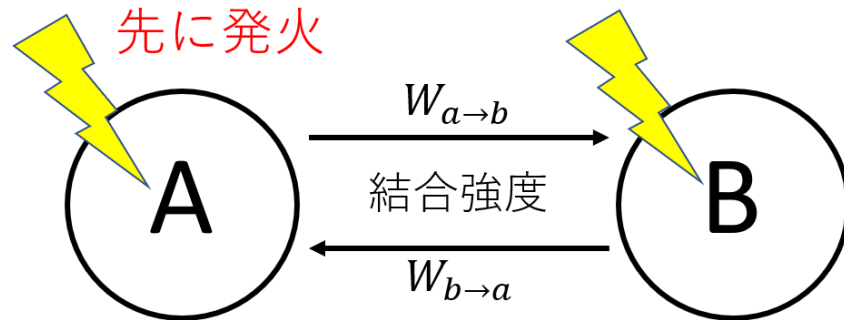


図 5.4-2 STDP の例

結合強度変化のスパイクタイミング依存性は、大脳皮質や海馬など、様々な脳の部位で観測されており、そのほとんどが図 5.4-3 のような関係性を示している。

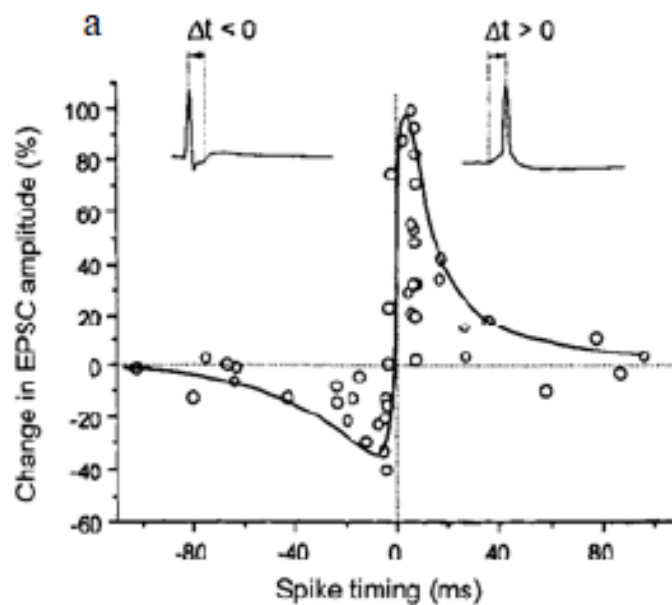


図 5.4-3 スパイクタイミングを変化させたときのシナプス後電位の変化[9]

この結果から、図 5.4-4 のように指数関数を用いてモデル化することができる。

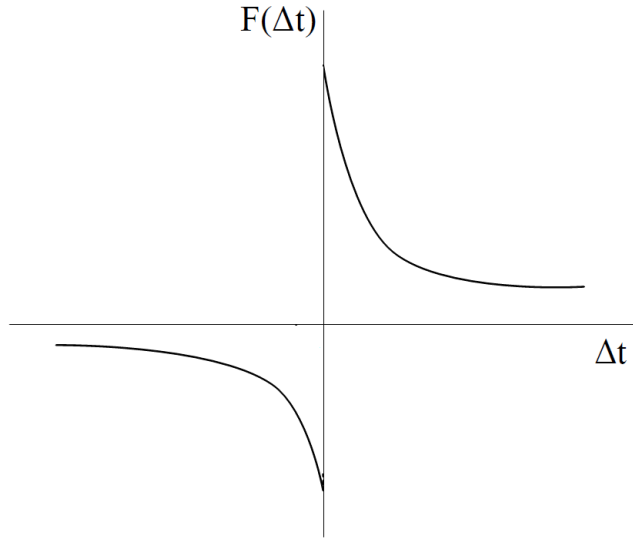


図 5.4-4 STDP の学習曲線  
F(Δt)は結合荷重の変化。

2つのニューロンについて着目すると、発火順序に従ってシナプス電流の送信側 (pre) と受信側 (post) が決まる。このとき、送信側の発火時間を $t_{pre}$ 、受信側の発火時間を $t_{post}$ とすると、 $\Delta t = t_{post} - t_{pre}$ と定義され、以下の式で値が更新される。

$\Delta t \geq 0$  のとき、

$$\Delta W = +A \exp\left(-\frac{|\Delta t|}{\tau_{STDP}}\right) \quad (15)$$

$\Delta t < 0$  のとき、

$$\Delta W = -A \exp\left(-\frac{|\Delta t|}{\tau_{STDP}}\right) \quad (16)$$

本モデルでは、同一音素内のアトラクタ間の結合を他のアトラクタ間の結合に対して少し深く成長させる必要があるため、通常の STDP 学習則を少し修正した。図 5.4-5 に示すように、 $\Delta t=0$  付近の $[-\varepsilon, \varepsilon]$ の領域で F(Δt)は一定の値をとるようにした。こうすることで、 $\Delta t=0$  付近でほぼ同時に発火したニューロン間の結合を成長させることができる。

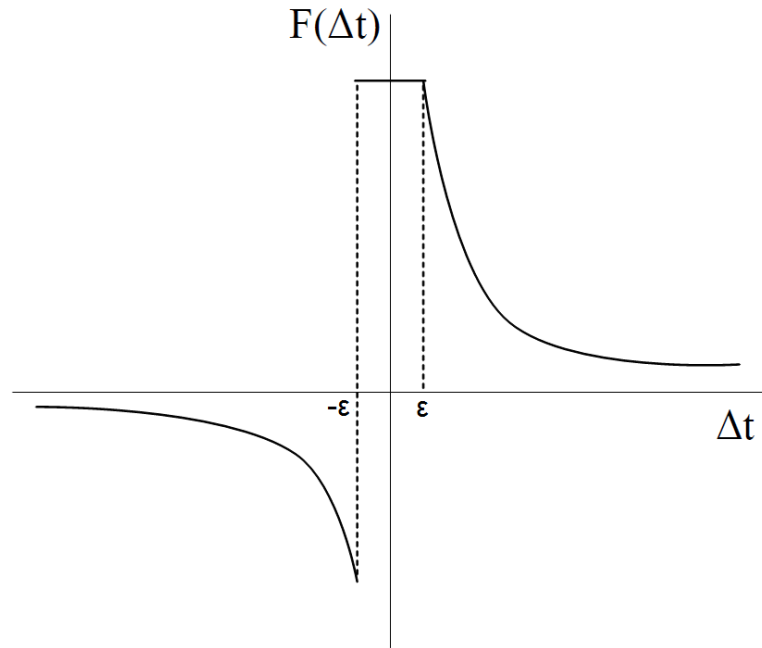


図 5.4-5 今回用いた学習曲線 [8]

$\Delta t > \varepsilon$  のとき、

$$\Delta W = +A \exp\left(-\frac{|\Delta t|}{\tau_{STDP}}\right) \quad (17)$$

$-\varepsilon \leq \Delta t \leq \varepsilon$  のとき、

$$\Delta W = +A \exp\left(-\frac{\varepsilon}{\tau_{STDP}}\right) \quad (18)$$

$\Delta t < -\varepsilon$  のとき、

$$\Delta W = -A \exp\left(-\frac{|\Delta t|}{\tau_{STDP}}\right) \quad (19)$$

FB 層で用いたパラメータを、表 5.4-1 に示す。

表 5.4-1 パラメータ

$W_{klmn}$	5.0	$V_{sp}$	50.0
$\tau_{LIF}$	30.0	$\theta$	8.0
$\tau_s$	30.0	$A$	0.01
$\lambda_s$	30.0	$W_{stdp}(\text{初期値})$	$1.0 \pm 0.5$

## 6. 結果

### 6.1 A1 層の単一シラブル音に対する応答

本研究では、図 2.2-1 のソナグラムの第二フォルマントのみを考え、以下のような単一シラブル音の刺激を神経回路モデルに対して入力した。それは、モデルを単純化するためとシラブル識別に関して第二フォルマントが主に寄与するためである。

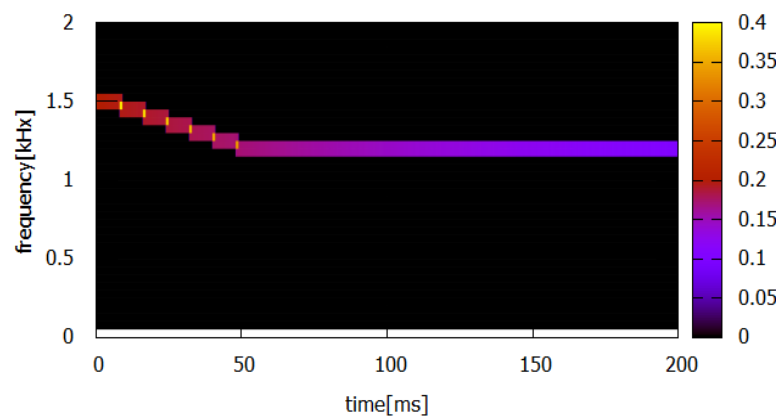


図 6.1-1 入力

このような入力に対する A1 層のあるユニットの挙動を図 6.1-2 に示す。まず、入力により興奮性ニューロンの膜電位が上昇する。興奮性ニューロンが活性化されると、続いて抑制性ニューロンの活性が上がり、興奮性ニューロンに作用することで興奮性ニューロンの活性が落ちる。このような変化を繰り返し、別のユニットへと興奮を伝えていく。

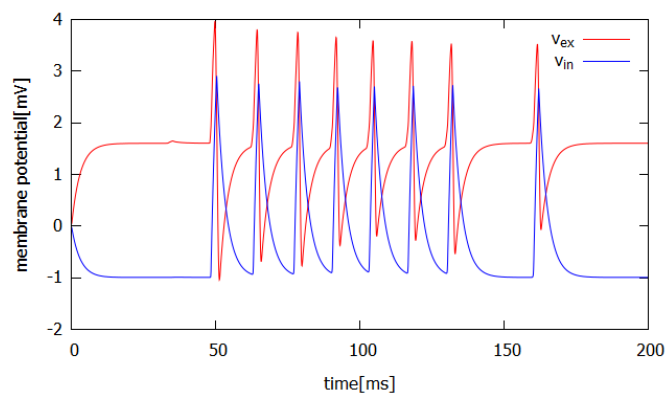


図 6.1-2 A1 層のあるユニットの膜電位変化  
赤：興奮性ニューロン、青：抑制性ニューロン

図 6.1-3 に A1 層全体の活動の時間的変化を示す。縦軸の周波数は、その周波数に反応するニューロン（興奮性ニューロン）の位置を表している。刺激の入力後、50 ms までは子音成分であり、周波数が変化するため、A1 層の応答は周波数軸を横切って活性が伝播していく。それ以降は周波数が変化しない母音成分であるため、周波数は変わらず、伝播軸に沿って活動が伝播していく。

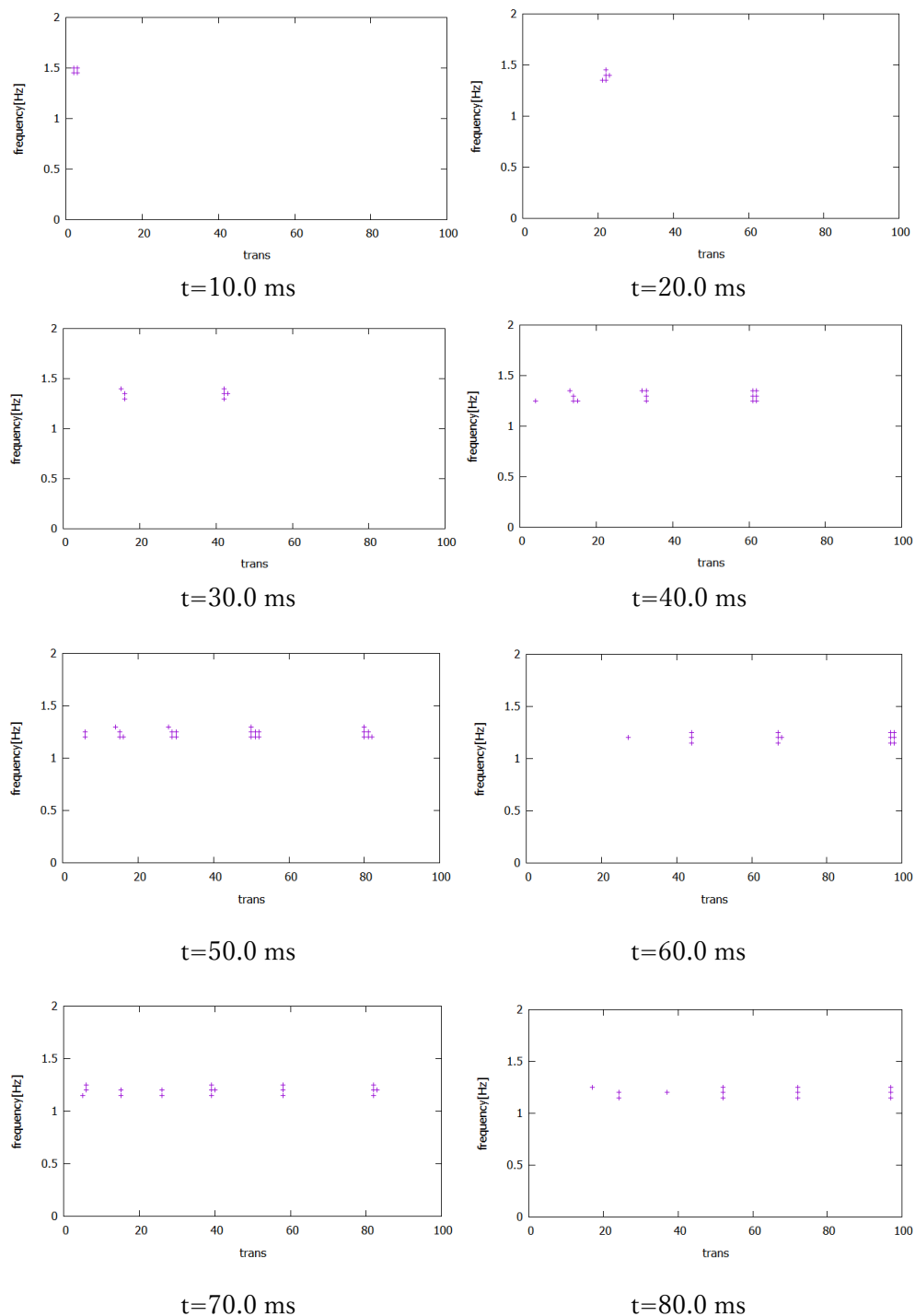


図 6.1-3 A1 層の応答

この結果は、Yamaguchi らの実験結果[4]と類似しており、このモデルにおいて一次聴覚野の興奮性と抑制性の神経活動の伝播を再現することができた。

## 6.2 FD 層のシラブル音に対する応答

一次聴覚野で得られた単一シラブル音に対する神経活動の時空間パターンから、FD 層で子音や母音の音の構成要素の特徴を抽出する。図 6.1-1 に、このシラブル刺激を何度も繰り返し与えて学習した特徴抽出細胞の応答結果を示す。

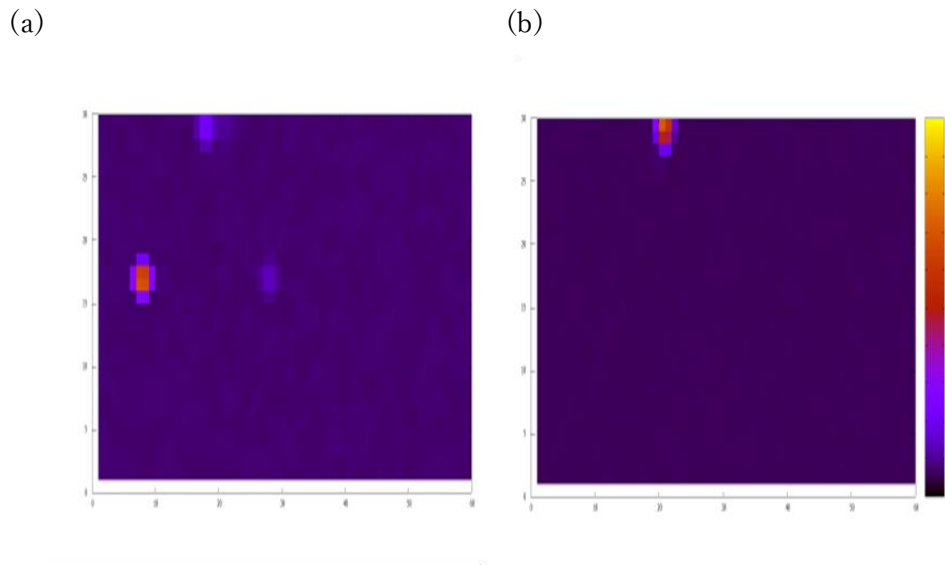


図 6.2-1 FD 層における特徴抽出細胞の活性

(a) 子音入力時 (b) 母音入力時

図 6.1-1a に示すように、子音は複数の周波数帯を横切って活性が伝播していくため、様々な場所に特徴抽出細胞ができる。それに対し、母音は伝播軸方向に沿ってのみ活性が移動していくため、周波数軸方向には移動しない細胞が応答した。

また、FD 層全体の出力  $\sum_{kl} P(x_{kl})$  は、次の図 6.2-2 のようになる。

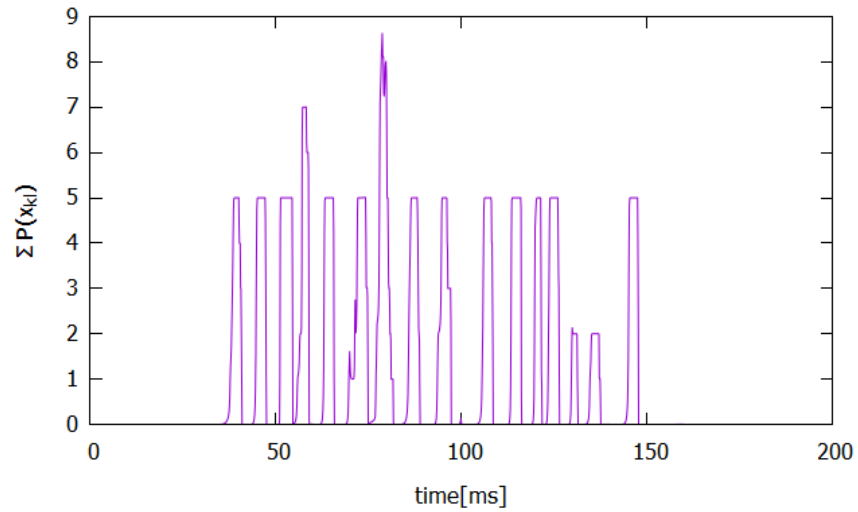


図 6.2-2 FD 層全体の出力

今回、図 6.1-1 の音情報を刺激として与えたため、50 ms までが子音成分、それ以降が母音成分である。しかし、図 6.2-2 のように、FD 層全体の出力として見ると、全体が連続的に発火し、子音情報と母音情報の区別がつかなくなる。そこで、理解できる音節速度のときには  $\beta$  帯域の活性が高くなるという Pefkou らの実験結果から、FD 層の各ニューロンに対して  $\beta$  帯域（14-21 Hz）の振動を加えた。8 種類の  $\beta$  周波数の振動を加えたときの FD 層の出力を以下に示す。

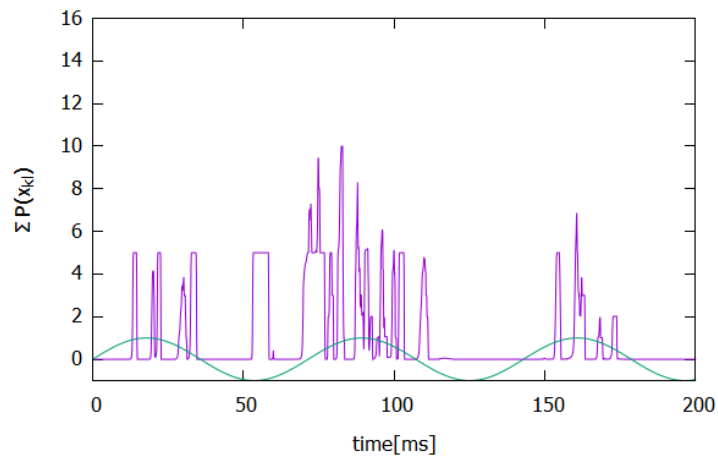


図 6.2-3 FD 層全体の出力（ $\beta$  振動：14 Hz）

緑： $\beta$  振動の正弦波。



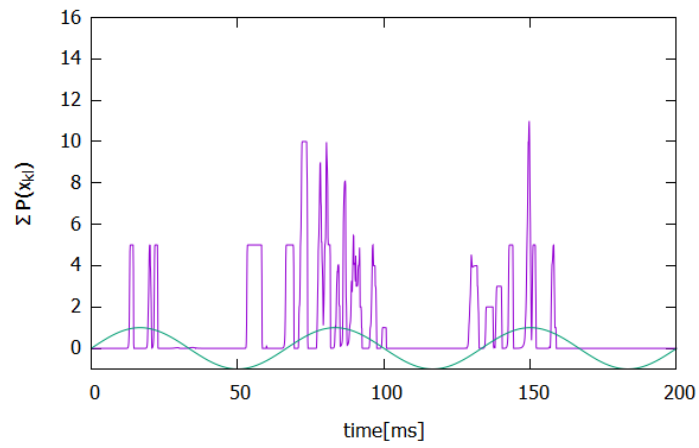


図 6.2-4 FD 層全体の出力 ( $\beta$  振動 : 15 Hz)

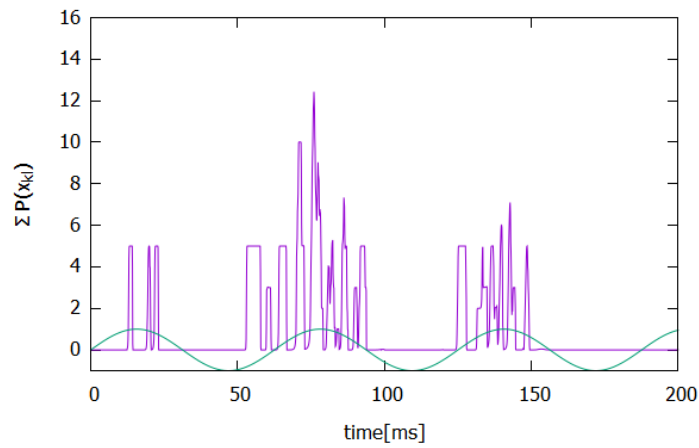


図 6.2-5 FD 層全体の出力 ( $\beta$  振動 : 16 Hz)

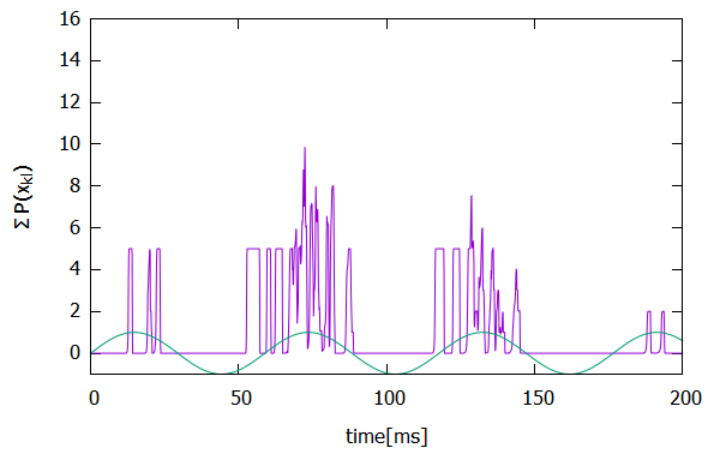


図 6.2-6 FD 層全体の出力 ( $\beta$  振動 : 17 Hz)

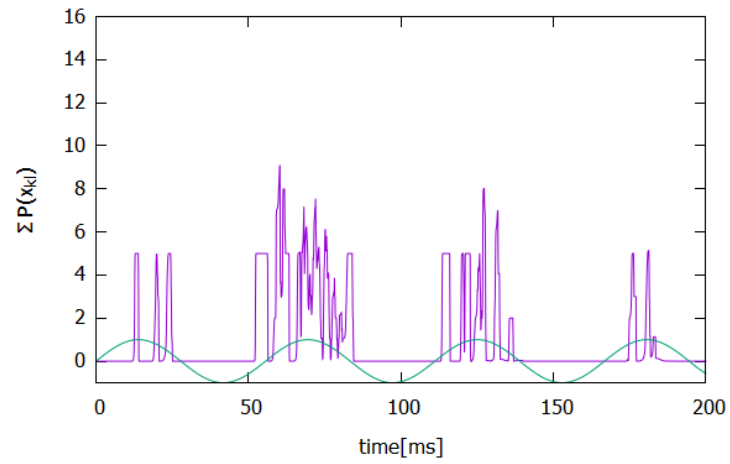


図 6.2-7 FD 層全体の出力 ( $\beta$  振動 : 18 Hz)

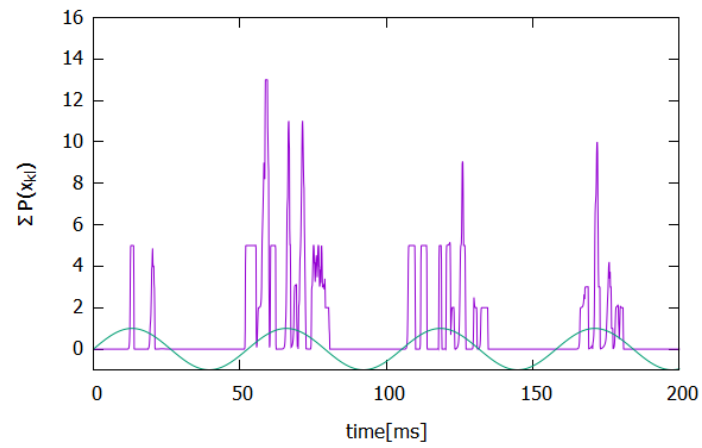


図 6.2-8 FD 層全体の出力 ( $\beta$  振動 : 19 Hz)

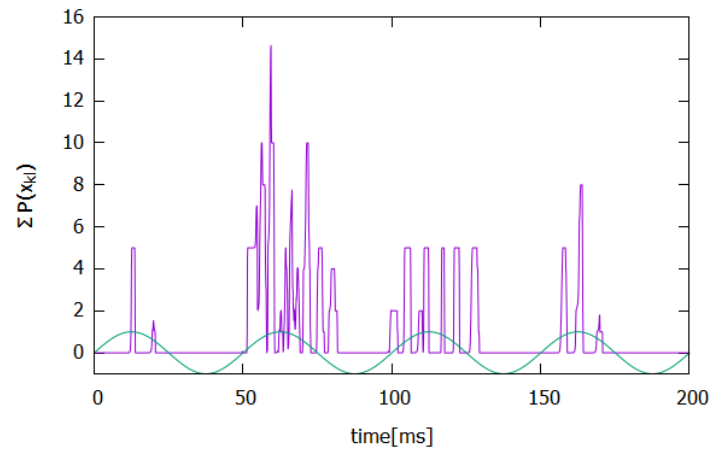


図 6.2-9 FD 層全体の出力 ( $\beta$  振動 : 20 Hz)

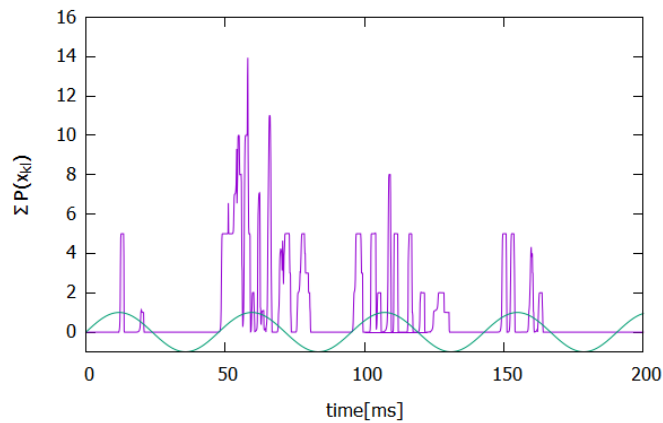


図 6.2-10 FD 層全体の出力 ( $\beta$  振動 : 21 Hz)

FD 層に振動を加えることで、出力を図 6.2-3~10 のように発火をいくつかのブロックにわけることができた。それぞれの発生している時間から、各ブロックを左から順に、子音 (C)、子音と母音の混在 (C+V)、母音 (V) (、母音) に対応する出力であると考えることができる。

次に、子音の持続時間と  $\beta$  帯域の関係性について調べた。図 6.1-1 において、子音部分は 50 ms であるが、子音の持続時間を変化させて、同様のシミュレーションを行った。以下に、 $\beta$  振動が 20 Hz のときに子音の長さを変化させた結果を示す。

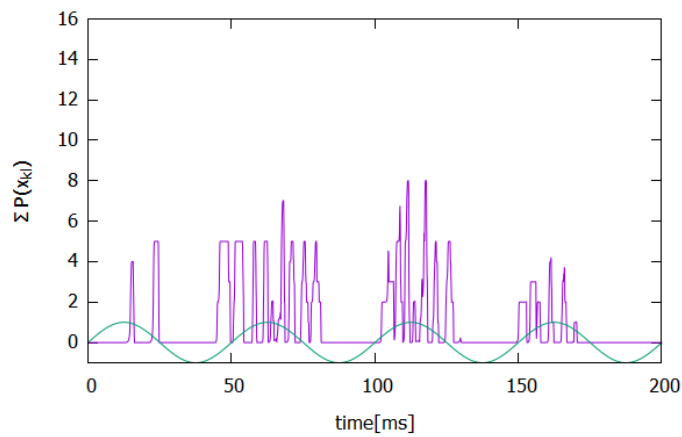


図 6.2-11 FD 層全体の出力 (子音 : 40 ms  $\beta$  振動 : 20 Hz)

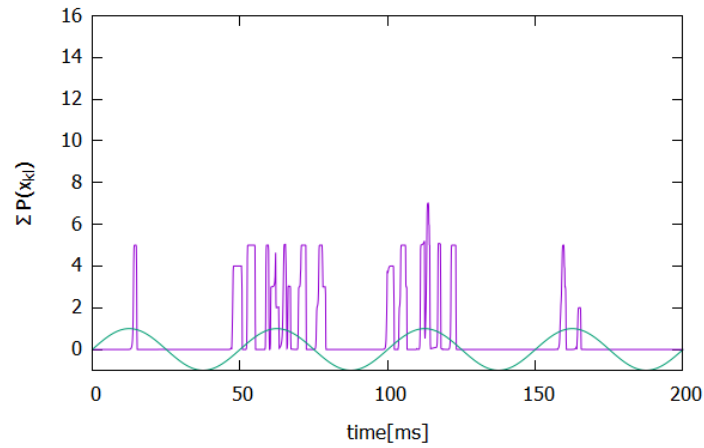


図 6.2-12 FD 層全体の出力（子音：60 ms  $\beta$  振動：20 Hz）

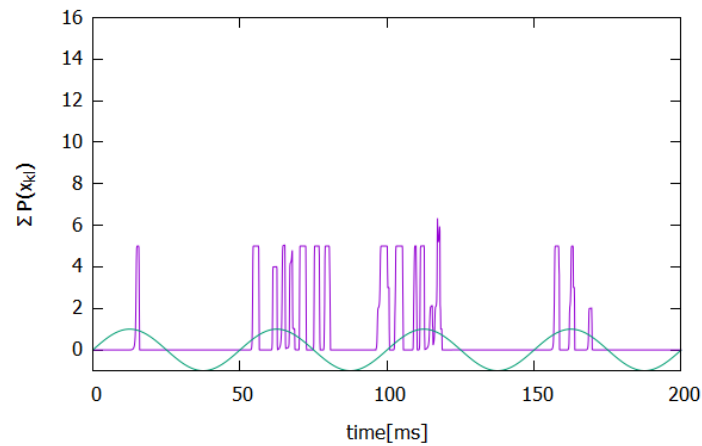


図 6.2-13 FD 層全体の出力（子音：70 ms  $\beta$  振動：20 Hz）

図 6.2-11~13 に示すように、子音の長さが変化しても  $\beta$  振動による子音/母音情報の分離能力は大きく影響を受けない。以上の結果から、 $\beta$  振動の役割として、音素の時間的に連続な入力を受けた際、子音や母音などの情報にセグメント化することが挙げられる。このように FD 層では、シラブルの構成要素である子音、母音やその接合部の情報を表現している。

また、実際の脳では、 $\beta$  リズムはある周波数帯域に分布していて、複数の周波数成分の合成波として作用している。そこで、各子音の長さにおいて、どの周波数が情報の分離に効果的に作用しているかを調べるため、FD 層の出力と各周波数の正弦波の相互相関（時間差 0 のピークの高さ）を計算した。その結果を図 6.2-14 に示す。

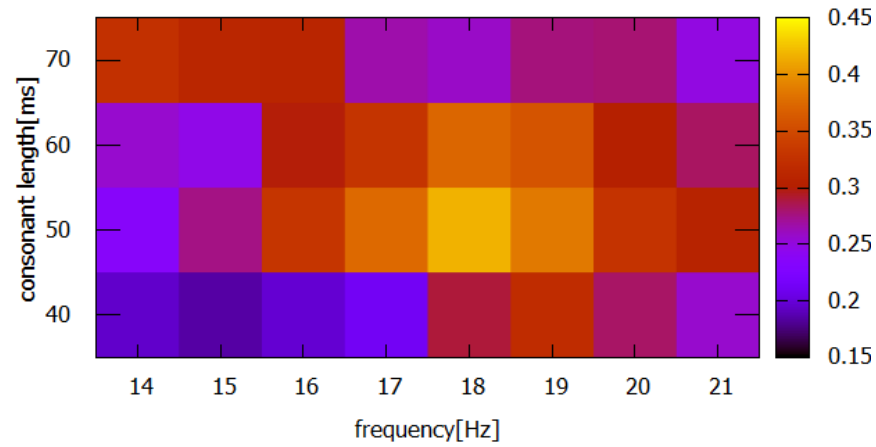


図 6.2-14 各子音の長さとは  $\beta$  帯域との相互相関

子音の長さが 50~60 ms のときは 18 Hz を中心に幅広い帯域との相互相関が高いことがわかる。逆に子音が短すぎたり、長すぎたりすると全体的に相関が低くなる。特に子音が短いときの結果は、Pefkou らが行った実験結果である、音の圧縮率が理解できなくなるほど高くなると  $\beta$  帯域の活性が落ちることと一致している [6]。

### 6.3 FB 層におけるシラブル内の子音—母音要素の結合

FB 層では、FD 層で分離した子音、母音要素を結合して一つのシラブル情報を表現する。FB 層の各ニューロンは、FD 層全体の出力  $\sum_{kl} P(x_{kl})$  を入力として受ける。図 6.3-1~6.3-3 に、図 6.1-1 のシラブル刺激に対する学習前と 2 つの学習時期のラスタプロットを示す。また、図 6.3-4 には、入力の想起期間での応答を示す。学習期間は 20000 ms、その後 800 ms 休止期間を挟んだ後、想起信号を送った。想起信号は学習信号の 33%を設定した。

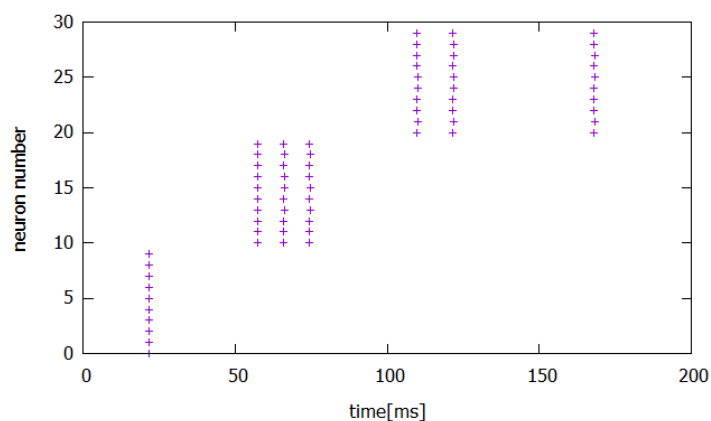


図 6.3-1 学習前の FB 層のラスタープロット  
ニューロン番号 0~9 が C、10~19 が C+V、  
20~29 が V に反応するニューロン。

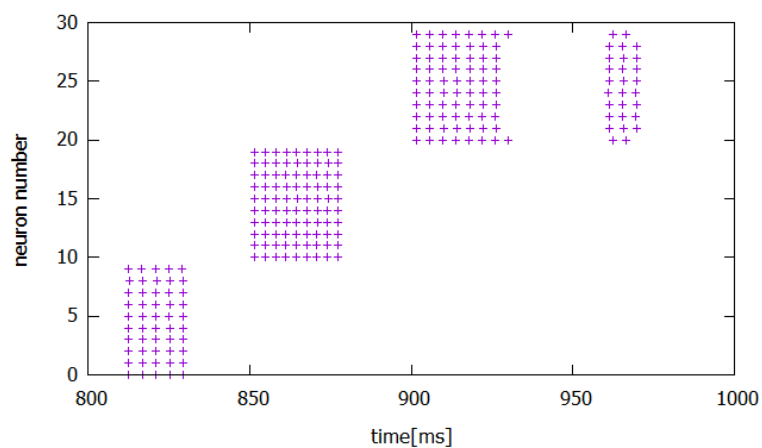


図 6.3-2 学習初期の FB 層のラスタープロット

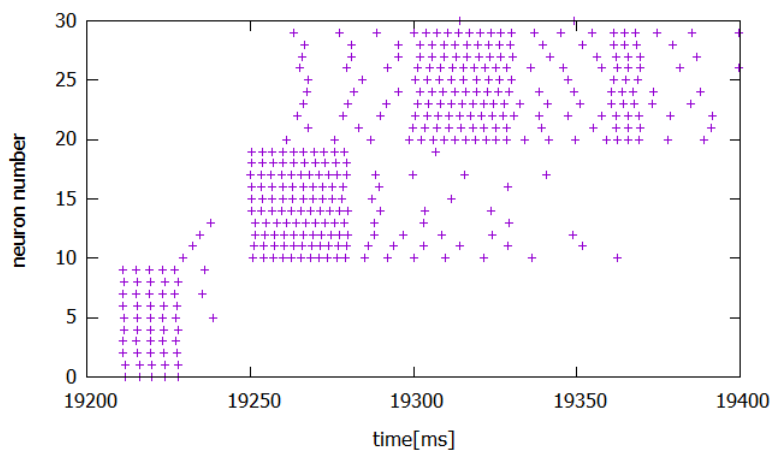


図 6.3-3 学習終期の FB 層のラスタープロット

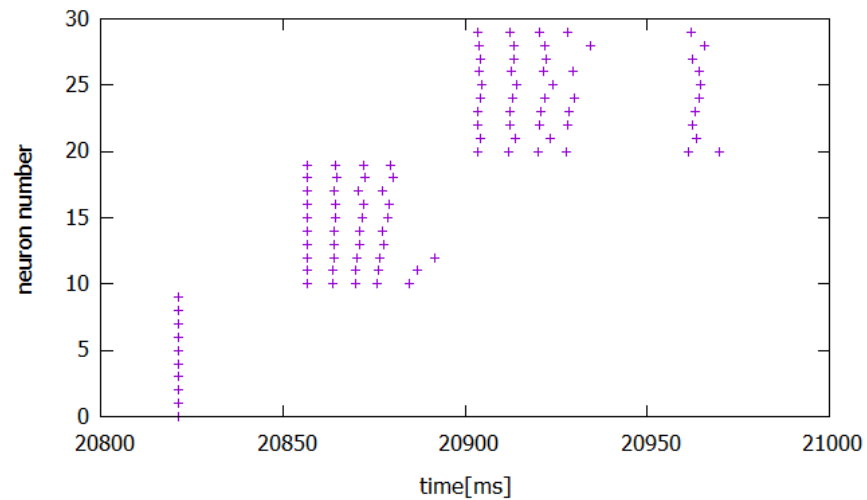


図 6.3-4 学習後、想起信号を入力したときのFB層のラスタプロット

図 6.3-1 と図 6.3-4 を比較すると、FB 層で分離した子音や母音をコードするニューロンの発火数が学習によって増加し、その関連がより強くなる。この結びつきの強化はひとまとまりのシラブルとしての認識を可能にする。また、学習終期になると母音の入力が入る前に母音を認識するニューロンがいくつか発火している。これは、STDP 学習により同一音節内の子音—母音の結合が強化されるからである。この結果により、音節情報は子音や母音の情報を保ちながらその相関を強くして一つのシラブルのアトラクタを形成する。

次に、FB 層内における  $\theta$  振動の効果について調べた。 $\theta$  レンジは 4-8 Hz であるが、今回は 5 Hz に固定してシミュレーションを行った。学習後の結果を以下の図 6.3-5 に示す。

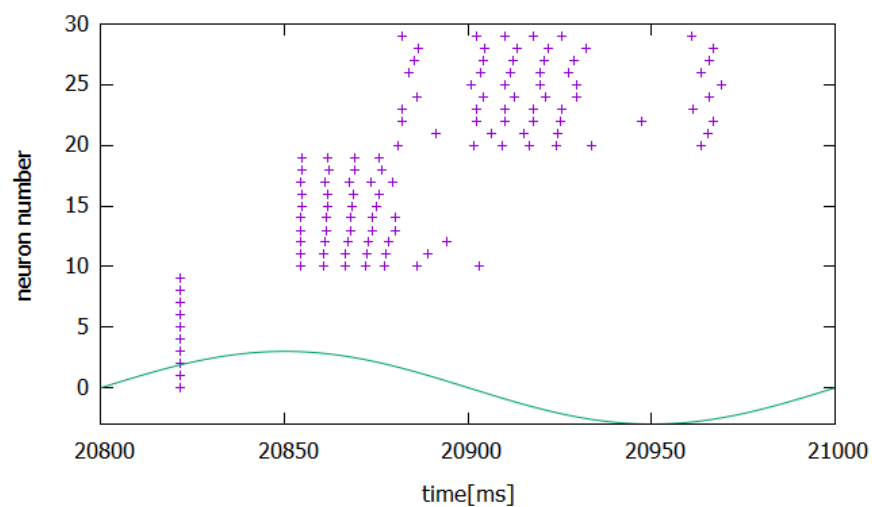


図 6.3-5 学習後のFB層のラスタプロット（ $\theta$  振動：5 Hz）緑： $\theta$  振動の正弦波

図 6.3-4 と図 6.3-5 を比べると、子音+母音のニューロン群と母音のニューロン群の発火数が増え、より強いアトラクタとして結合されていることがわかる。また、それぞれのニューロン間の結合強度は以下のようにになっている。

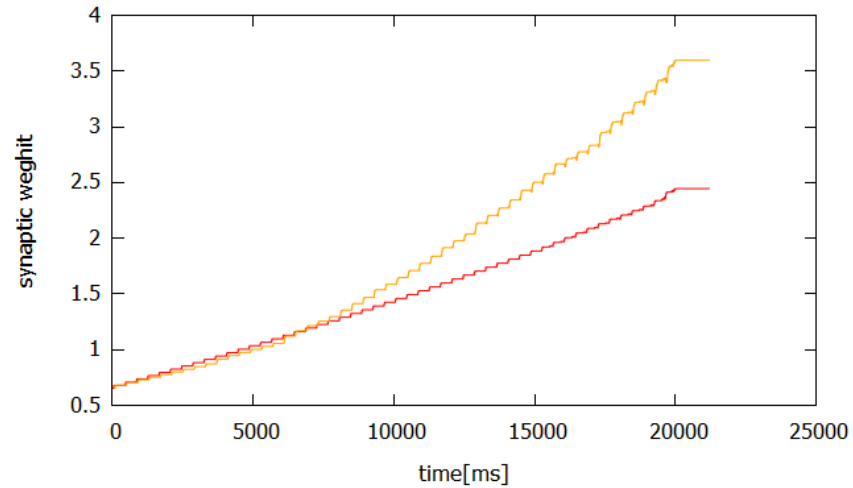


図 6.3-6 各ニューロン間の結合強度の変化 ( $\theta$  振動なし)

赤：C→C+V への結合強度

橙：C+V→V への結合強度

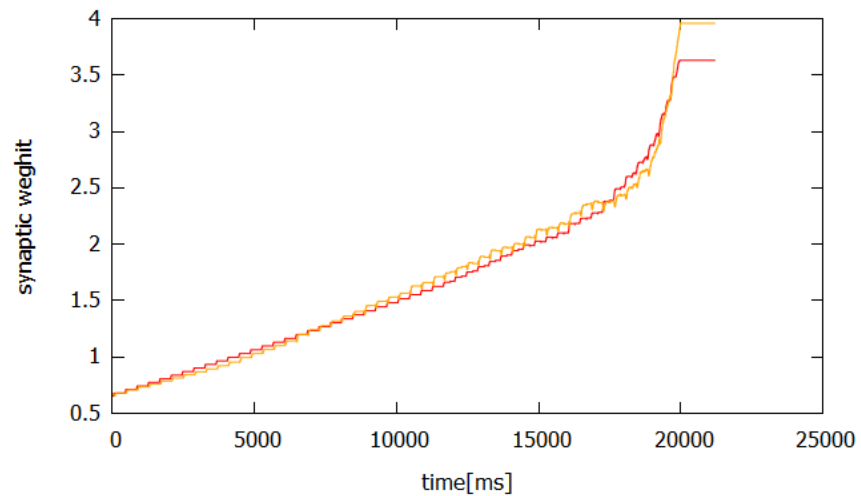


図 6.3-7 各ニューロン間の結合強度の変化 ( $\theta$  振動あり)

赤：C→C+V への結合強度

橙：C+V→V への結合強度

図 6.3-6 と図 6.3-7 を比較すると、 $\theta$  振動があると、無い場合よりも学習後の層内結合強度が大きくなっていることがわかる。これらの結果から、 $\theta$  振動は同一音節内の子音と母音の結合をより強くし、音節としての認知に貢献していると考えられる。



## 6.4 連続したシラブルに対する各部位の応答

ここまで、図 6.1-1 のような単一の音素刺激に対する各層の応答を見てきたが、現実の世界において単一の音素（単一シラブル）のみを聞き取る機会はほとんどなく、連続した音素を聞くことがほとんどである。そこで、図 6.4-1 のように、2つの音素を連続で与えた。

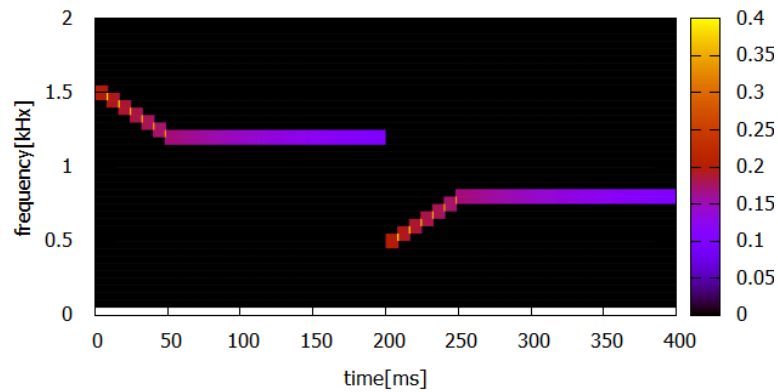


図 6.4-1 連続音素の入力

このような刺激を与えたとき、シラブルが切り替わる場所の A1 層の応答を以下に示す。

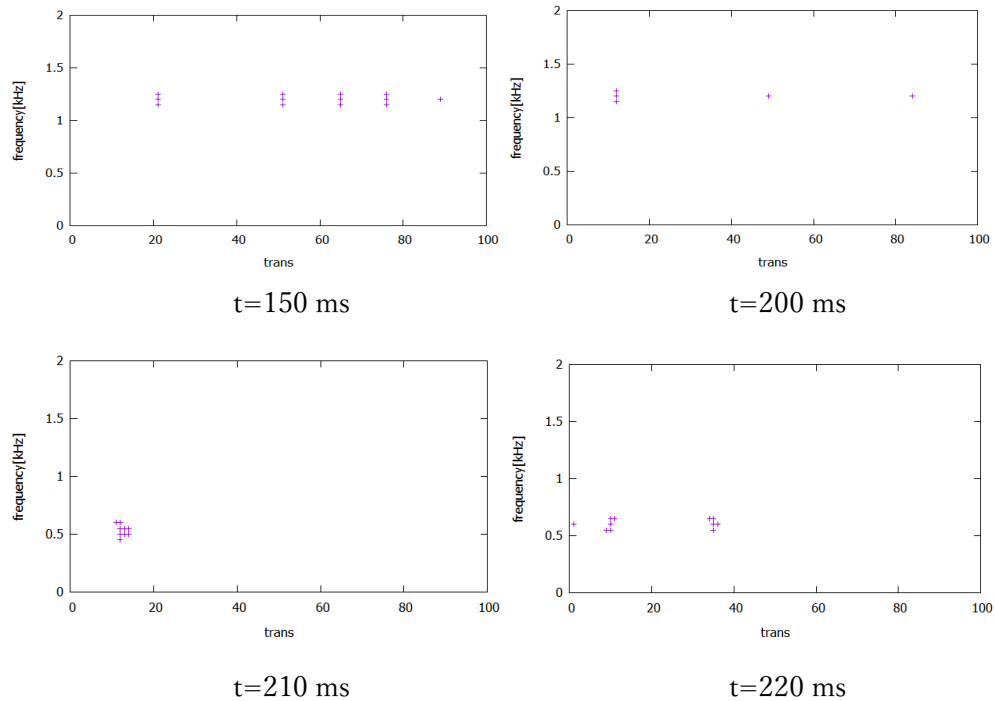


図 6.4-2 連続のシラブルに対する A1 層の応答

図 6.4-2 のように、連続したシラブルを与えた場合、それぞれ個別で刺激を与えた場合と同じような応答を示す。これは、2 つめのシラブルの子音によって A1 層のユニットが活性化する前に、1 つめのシラブルの母音の活性が落ちるためである。

また、FD 層の出力は次のようになる。

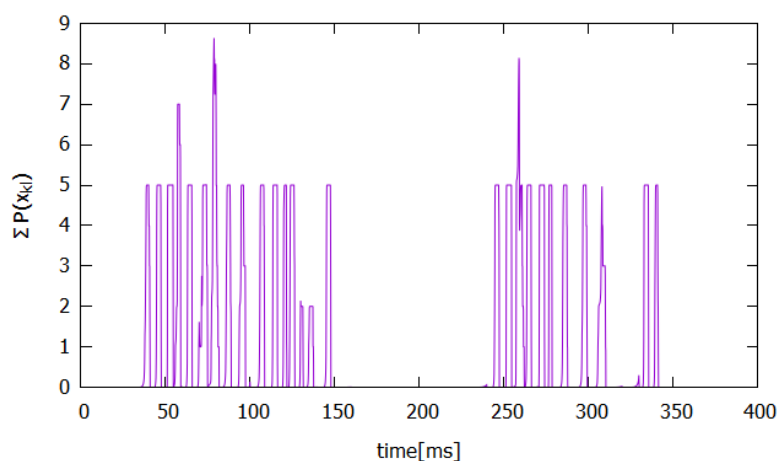


図 6.4-3 FD 層の出力 ( $\beta$  振動なし)

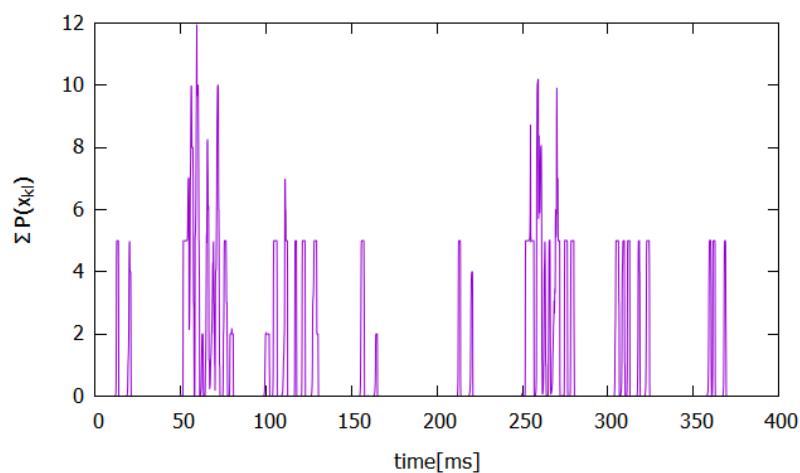


図 6.4-4 FD 層の出力 ( $\beta$  振動: 20 Hz)

このように、音素が連続で入力された場合であっても、単一音素のときと同様に、 $\beta$  振動を用いることで情報の分離をすることができる。

図 6.4-4 の出力に対する FB 層の応答（ラスタプロット）を以下に示す。

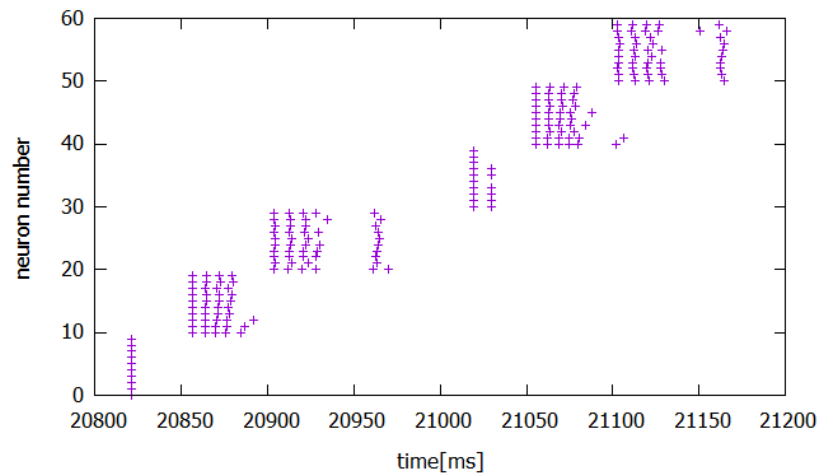


図 6.4-5 FB 層のラスタプロット

ニューロン番号 0~29 音素 1 (C1、C1+V1、V1) に、  
ニューロン番号 30~59 が音素 2 に対応するニューロン。

図 6.4-5 より、連続で音素が入力された場合でも、それぞれの音素に対して単一音素のときと同様の結果が得られた。

以上の結果より、本モデルでは刺激のシラブルが単一か連続かに関係なく、同様の結果を得ることができる。

## 7. 結論

本研究では、一次聴覚野にて生じる神経活動の時空間パターンから、音情報を階層的に処理するモデルを提案した。A1層では、Yamaguchiらの実験結果をもとに、一次聴覚野にて活性が伝播していく様子を再現した。FD層では、A1層の時空間パターンから子音や母音などの音の要素的な特徴をKohonen mapを用いることによって抽出した。FB層では、FD層によって得られた特徴をまとめあげ、ひとつの音素（シラブル）としてアトラクタを形成し、表現することができた。

また、脳内リズムの効果として、 $\beta$ 振動はA1における連続的な入力から音素内の子音や母音の情報を分離することができる。連続した音情報をセグメント化することで、音素の認知に貢献していると考えられる。 $\theta$ 振動はひとつの音素間の結合を強化し、さらに大きなまとまりであるシラブルとしての認識に貢献している。

## 8. 今後の課題

本研究では、音節情報として図6.1-1のような刺激を用いたが、音声は実際にはいくつかのフォルマントで構成されている。複数のフォルマントも入力すれば、より自然音声に近い刺激でシミュレーションを行うことができると考えられる。また、脳内リズムについても、 $\beta$ 、 $\theta$ などは幅広い周波数帯を持つが、認知に必要な周波数をどのように決定しているのか、他の周波数帯の脳内リズム（ $\alpha$ や $\gamma$ など）はどう関与しているかなど、考えられることは多くある。

## 9. 謝辞

本研究を進めるにあたり、様々なご指導を頂きました主任指導教員の檜森与志喜教授に感謝致します。また、日頃より研究活動だけでなく、学生生活においても様々なアドバイスをしていただいた研究室の先輩、同期、後輩の皆様に感謝致します。

## 10. 参考文献

- [1] 入来正躬, 外山敬介, “生理学 I”, 文光堂(1986)
- [2] 塚原仲晃, “脳の情報処理”, 朝倉書店(1984)
- [3] Y.Yamaguchi, J.Horikawa, and I.Taniguchi, “Neural Dynamics of Vocal Processing in the Auditory Cortex”, *Biophysical Neural Networks*, 343-362, (2001)
- [4] 山口陽子, 谷口郁雄, “聴覚皮質神経回路の波動伝播活動”, *信学技報 NC95-142*, 198-204(1996)
- [5] Edward F Chag, Jochem W Rieger, Keith Johnson, Mitchel S Berger, Nicholas M Barbaro and Robert T Knight, "Categorical speech representation in human superior temporal gyrus", *Nature neuroscience*, 1428-1432(2010)
- [6] Maria Pefkou, Luc H. Arnal, Lorenzo Fontolan, and Anne-Lise Giraud, “ $\theta$  -Band and  $\beta$  -Band Neural Activity Reflects Independent Syllable Tracking and Comprehension of Time-Compressed Speech”, *The Journal of Neuroscience*, 37(33):7930-7938(2017)
- [7] Ann-Lise Giraud, David Poeppel, “Cortical oscillations and speech processing: emerging computational principles and operations”, *Nature neuroscience*, 15(4),511-517(2012)
- [8] Kazuhisa Fujita, Yusuke Hara, Youichi Suzukawa, Yoshiki Kashimori, “Decoding Word Information from Spatiotemporal Activity of Sensory Neurons”. *Cogn. Comput.* 6:145–157(2014)
- [9] Guo-qiang Bi, Mu-ming Poo, "Synaptic Modifications in Cultured Hippocampal Neurons: Dependence on Spike Timing, Synaptic Strength, and Postsynaptic Cell Type", *The Journal of Neuroscience*, 18(24):10464–10472(1998)

## 11. 付録

FB 層の結果について、様々な  $\theta$  振動数を用い、パラメータを調整して追加計算した結果を以下に示す。

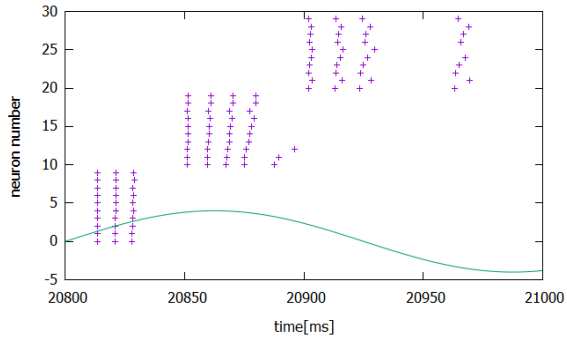


図 1. 4 Hz

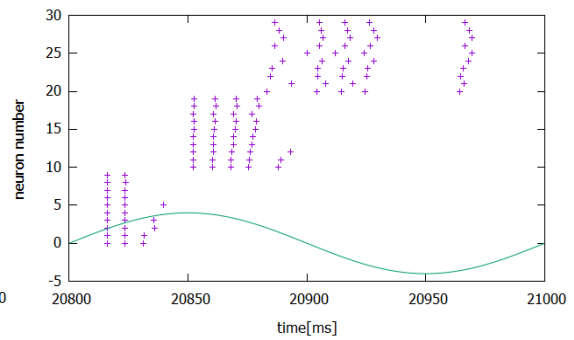


図 2. 5 Hz

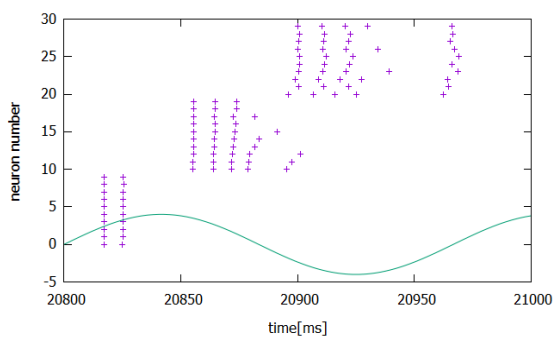


図 3. 6 Hz

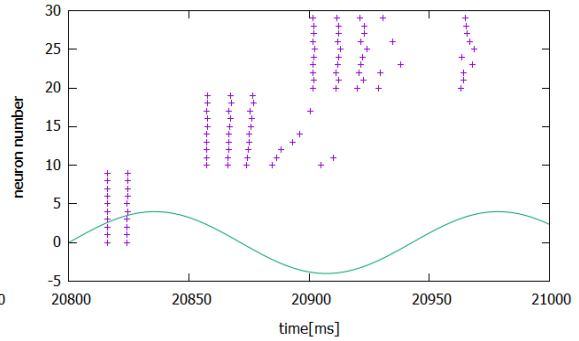


図 4. 7 Hz

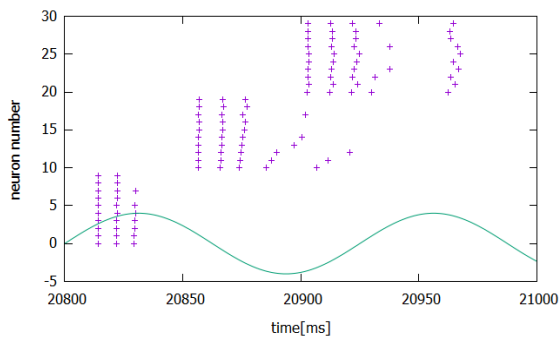


図 5. 8 Hz

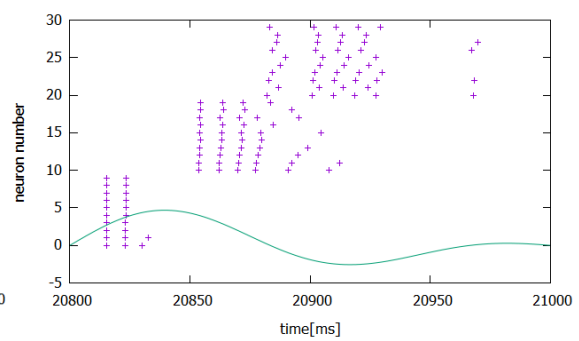


図 6. 4-8 Hz の合成波

高周波数のとき、C+V の発火が若干長く持続する。これは、学習時に正弦波の山が二回来る（山するときニューロンが発火しやすくなる）ため、結合が育ちやすいからだと思われる。合成波の場合でも、単一周波数のときと同様の結果が得られる。